

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:47:31; Search time 50.1771 Seconds

(without alignments)
371.762 Million cell updates/sec

Title: US-10-054-873-5

Perfect score: 294

Sequence: 1 FVNHGCGSHLVEALYLIVCG.....IVDQCTSIGSLVQLNYCN 52

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_23sep04:*
1: geneseq1980s:*
2: geneseq2000s:*
3: geneseq2001s:*
4: geneseq2002s:*
5: geneseq2003as:*
6: geneseq2003bs:*
7: geneseq2004s:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	52	2	AA42859 Human ins
2	294	100.0	56	2	AA68901 Human pro
3	294	100.0	56	2	AA78665 Proinsulin
4	294	100.0	63	2	AA68900 Human pro
5	294	100.0	96	2	AA68899 Human pro
6	294	100.0	96	2	AA78662 Fusion pr
7	294	100.0	107	2	AA42860 hGH-mini-
8	294	100.0	116	2	AA98897 SOD-proin
9	294	100.0	137	2	AA71692 Marling fa
10	294	100.0	145	2	AA71694 Marling fa
11	294	100.0	146	2	AA71695 Marling fa
12	294	100.0	150	2	AA42861 Chimeric
13	294	99.0	57	2	AA04582 Proinsulin
14	287	97.6	52	2	AA04582 Proinsulin
15	283.5	96.4	53	2	AA65883 DI-Arg-(B
16	283.5	96.4	53	2	AA65883 DI-Arg-(B
17	283.5	96.4	160	2	AA79056 Glycosp
18	282.5	96.1	117	2	AA98896 SOD-proin
19	281.5	95.7	60	1	AA20002 Human pro
20	281	95.6	58	7	ADP15278 Human al
21	281	95.6	58	7	ADP15279 Human al
22	281	95.6	58	7	ADP15279 Human al
23	281	95.6	58	7	ADP15279 Human al
24	281	95.6	58	7	ADP15279 Human al
25	281	95.6	58	7	ADP15279 Human al

26	281	95.6	662	7	ADP16456 Human al
27	281	95.6	662	7	ADP16456 Human al
28	281	95.6	667	7	ADP15064 Human al
29	281	95.6	667	7	ADP15065 Human al
30	281	95.6	667	7	ADP15065 Human al
31	281	95.6	667	7	ADP15065 Human al
32	279	94.9	62	7	ADP16653 Human al
33	279	94.9	62	7	ADP16606 Human al
34	279	94.9	62	7	ADP16607 Human al
35	279	94.9	62	7	ADP16607 Human al
36	279	94.9	62	7	ADP16607 Human al
37	279	94.9	62	7	ADP16607 Human al
38	279	94.9	62	7	ADP16607 Human al
39	279	94.9	62	7	ADP16607 Human al
40	279	94.9	62	7	ADP16607 Human al
41	279	94.9	62	7	ADP16607 Human al
42	279	94.9	62	7	ADP16607 Human al
43	279	94.9	62	7	ADP16607 Human al
44	278.5	94.7	51	3	AA012771 Human mat
45	278.5	94.7	51	3	AA012771 Human mat

ALIGNMENTS

RESULT 1
AA42859
ID AA42859 standard; protein; 52 AA.

AA42859;
19-JAN-2000 (first entry)

Human insulin precursor, SEQ ID 5.

Insulin; precursor; growth hormone; chaperone; intramolecular; folding; conformation; chimeric protein; cleavable; recombinant; production; yield.

OS Homo sapiens.

PN WO9950302-A1.

PD 07-OCT-1999.

PF 31-MAR-1998; 98WO-CN000052.

PR 31-MAR-1998; 98WO-CN000052.

(TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

Gan Z;

WPI; 1999-610839/52.

New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.

Claim 12; Page 29-30; 46pp; English.

This sequence represents a human insulin precursor comprising insulin A and B chains. This insulin precursor is a component of the chimeric proteins hGH-mini-proinsulin (AA42860) and the chimeric protein given in AA42861. These chimeric proteins additionally contain an N-terminal fragment of human growth hormone (hGH) and a cleavable peptide linker (AA42857). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only

CC protect insulin sequences from intracellular degradation by a
 CC microorganism host, but also promote the folding of the fused insulin
 CC precursor, facilitate the solubility of the fusion protein and decrease
 CC the intermolecular interactions among the fusion proteins, thus allowing
 CC folding of the fused insulin precursor at commercially useful high
 CC concentrations. The procedural steps of cyanogen bromide cleavage,
 CC oxidative sulfitolysis and related purification steps can thus be
 CC eliminated, along with the use of high concentrations of mercaptan or the
 CC use of hydrophobic adsorbent resins

XX Sequence 52 AA;

Query Match 100.0%; Score 294; DB 2; Length 52;

Best Local Similarity 100.0%; Pred. No. 1.4e-26; Mismatches 0; Indels 0; Gaps 0;

Db 1 FVNHLCGSHLVEALYVGSGRGPFPKTRGIVECCCTSGSLYLENYCN 52
 1 FVNHLCGSHLVEALYVGSGRGPFPKTRGIVECCCTSGSLYLENYCN 52

RESULT 2

ID AAR68901 standard; peptide; 56 AA.

AC AAR68901;

DT 25-MAR-2003 (revised)

DT 02-MAR-1995 (first entry)

XX Human pro-insulin 3.

XX Pro-insulin; A-chain; B-chain; C-chain; disulphide; mercaptan;

XX chaotropic agent.

XX Homo sapiens.

XX EP600372-A1.

XX 08-JUN-1994.

XX 25-NOV-1993; 93EP-00116993.

XX 02-DEC-1992; 92DE-04240420.

XX (PARH) HOECHST AG.

XX Obermeyer R, Gerl M, Ludwig J, Sabel W;

XX WPI; 1994-17718/22.

PT Prodn. of pro-insulin with correct di:disulphide bridges - by treating
 PT recombinant precursor protein with mercaptan in alkali and in presence of
 PT chaotropic agent, then isolation on hydrophobic resin.

PS Disclosure; Page 12; 15pp; German.

XX pro-insulin is produced by treating recombinant precursor protein with a
 CC mercaptan to provide 2-10 SH residues per Cys residue, in presence of a
 CC chaotropic agent and in aq. medium of pH 10-11, treating the prod. with 3
 CC -50 g hydrophobic adsorbent resin per 1 eq. medium of pH 4-7, isolating
 CC the adsorbed resin and pro-insulin and desorbing the pro-insulin. This
 CC method produces pro-insulin with correctly bonded Cys bridges. Compared
 CC with known methods it involves fewer stages (esp. no sulfitolysis or
 CC cyanogen bromide cleavage) and overall losses during purification are
 CC reduced, i.e. the process is quicker and gives better yields. Sequences
 CC of insulin chain A, B and C are given in AAR6895-97. Sequences of pro-
 CC insulin 1-4 are given in AAR6898-901. (Updated on 25-MAR-2003 to correct
 CC PN field.)

XX Sequence 56 AA;

Query Match 100.0%; Score 294; DB 2; Length 56;

Best Local Similarity 100.0%; Pred. No. 1.5e-26; Mismatches 0; Indels 0; Gaps 0;

Db 1 FVNHLCGSHLVEALYVGSGRGPFPKTRGIVECCCTSGSLYLENYCN 52
 5 FVNHLCGSHLVEALYVGSGRGPFPKTRGIVECCCTSGSLYLENYCN 56

RESULT 3

ID AAR78665 standard; protein; 56 AA.

AC AAR78665;

DT 03-APR-1996 (first entry)

XX Proinsulin sequence 3.

XX Proinsulin; post-translational modification; recombinant production;

XX protein folding; conformation.

XX Synthetic.

XX Key Location/Qualifiers

XX Region

XX /label= R2

XX /note= "a peptide of 4 amino acids"

XX Peptide

XX /label= R1-(B2-B29)-Y

XX /note= "human insulin B-chain"

XX Region

XX /label= X

XX /label= Gly-(A2-A20)-R3

XX /note= "human insulin A-chain"

XX EP600372-A2.

XX 23-AUG-1995.

XX 09-FEB-1995; 95EP-00101748.

XX 18-FEB-1994; 94DE-04405179.

XX (PARH) HOECHST AG.

XX Obermeyer R, Gerl M, Ludwig J, Sabel W;

XX WPI; 1995-284754/38.

PT Isolation of insulin that is correctly post-translationally processed -
 PT by reacting pro-insulin with a mercaptan in the presence of a chaotropic
 PT agent and purification, after absorption to hydrophobic resin.

PS Example 2; Page 13; 16pp; German.

XX The present sequence is an example of a proinsulin molecule corresp. to
 CC the general formula R2-R1-(B2-B29)-Y-X-Gly-(A2-A20)-R3 (II). In formula
 CC (II): X = Lys; Arg or a peptide of 2-35 amino acids contg. Lys or Arg at
 CC the N- and C-termini; Y = a natural amino acid; R1 = Phe or a bond; R2 =
 CC H, Arg, Lys, a peptide of 2-45 amino acids contg. Arg or Lys at the N-
 CC and C-termini; R3 = a natural amino acid; (A2-A20) and (B2-B29) are the
 CC insulin A- and B-chain sequences from human or other insulin. The
 CC proinsulin molecule (produced in recombinant E. coli) is reacted with
 CC mercaptan at a ratio of 2-10 SH residues of mercaptan per Cys residue of
 CC proinsulin. The reaction takes place in the presence of a chaotropic
 CC auxiliary agent at pH 10-11 and results in proinsulin with correctly
 CC linked cystine bridges. Reaction with trypsin and opt. carboxypeptidase B
 CC yields correctly folded insulin. The insulin is isolated by absorption on
 CC a hydrophobic resin

XX Sequence 56 AA;

Query Match 100.0%; Score 294; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 1.5e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYIVCGERGFYPTKRGIVGQCTTCSLYOLENYCN 52
 DB 5 FVNOHLCGSHLVEALYIVCGERGFYPTKRGIVGQCTTCSLYOLENYCN 56

RESULT 4
 AAR68900
 ID AAR68900 standard; peptide; 63 AA.
 XX
 AC AAR68900;
 XX

DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)

DE Human pro-insulin 4.

KM Pro-insulin; A-chain; B-chain; C-chain; disulphide; mercaptan;
 KW chaotropic agent.

OS Homo sapiens.

PN EP600372-A1.

PD 08-JUN-1994.

PF 25-NOV-1993; 93EP-00118993.

PR 02-DEC-1992; 92DE-04240420.

PA (FARH) HOECHST AG.

PI Obermeier R, Gerl M, Ludwig J, Sabel W;

DR WPI; 1994-177718/22.

PT Prodn. of pro-insulin with correct disulphide bridges - by treating
 PT recombinant precursor protein with mercaptan in alkali and in presence of
 PT chaotropic agent, then isolation on hydrophobic resin.

PS Disclosure; Page 11-12; 15pp; German.

CC Pro-insulin is produced by treating recombinant precursor protein with a
 CC mercaptan to provide 2-10 SH residues per Cys residue, in presence of a
 CC chaotropic agent and in aq. medium of pH 10-11, treating the prod. with 3
 CC -50 g hydrophobic adsorbent resin per 1 aq. medium of pH 4-7, isolating
 CC the adsorbed resin and pro-insulin and desorbing the pro-insulin. This
 CC method produces pro-insulin with correctly bonded Cys bridges. Compared
 CC with known methods it involves fewer stages (esp. no sulphydolytic or
 CC cyanogen bromide cleavage) and overall losses during purification are
 CC reduced, i.e. the process is quicker and gives better yields. Sequences
 CC of insulin chain A, B and C are given in AAR68895-97. Sequences of pro-
 CC insulin 1-4 are given in AAR68898-901. (Updated on 25-MAR-2003 to correct
 CC PN field.)

XX SQ Sequence 63 AA;

Query Match 100.0%; Score 294; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.7e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYIVCGERGFYPTKRGIVGQCTTCSLYOLENYCN 52
 DB 12 FVNOHLCGSHLVEALYIVCGERGFYPTKRGIVGQCTTCSLYOLENYCN 63

RESULT 5
 AAR68899
 ID AAR68899 standard; peptide; 96 AA.
 XX

AC AAR68899;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)

DE Human pro-insulin 2.

KM Pro-insulin; A-chain; B-chain; C-chain; disulphide; mercaptan;
 KW chaotropic agent.

OS Homo sapiens.

PN EP600372-A1.

PD 08-JUN-1994.

PF 25-NOV-1993; 93EP-00118993.

PR 02-DEC-1992; 92DE-04240420.

PA (FARH) HOECHST AG.

PI Obermeier R, Gerl M, Ludwig J, Sabel W;

DR WPI; 1994-177718/22.

PT Prodn. of pro-insulin with correct disulphide bridges - by treating
 PT recombinant precursor protein with mercaptan in alkali and in presence of
 PT chaotropic agent, then isolation on hydrophobic resin.

PS Disclosure; Page 11; 15pp; German.

CC Pro-insulin is produced by treating recombinant precursor protein with a
 CC mercaptan to provide 2-10 SH residues per Cys residue, in presence of a
 CC chaotropic agent and in aq. medium of pH 10-11, treating the prod. with 3
 CC -50 g hydrophobic adsorbent resin per 1 aq. medium of pH 4-7, isolating
 CC the adsorbed resin and pro-insulin and desorbing the pro-insulin. This
 CC method produces pro-insulin with correctly bonded Cys bridges. Compared
 CC with known methods it involves fewer stages (esp. no sulphydolytic or
 CC cyanogen bromide cleavage) and overall losses during purification are
 CC reduced, i.e. the process is quicker and gives better yields. Sequences
 CC of insulin chain A, B and C are given in AAR68895-97. Sequences of pro-
 CC insulin 1-4 are given in AAR68898-901. (Updated on 25-MAR-2003 to correct
 CC PN field.)

XX SQ Sequence 96 AA;

Query Match 100.0%; Score 294; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.5e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYIVCGERGFYPTKRGIVGQCTTCSLYOLENYCN 52
 DB 45 FVNOHLCGSHLVEALYIVCGERGFYPTKRGIVGQCTTCSLYOLENYCN 96

RESULT 6
 AAR78662
 ID AAR78662 standard; protein; 96 AA.
 XX
 AC AAR78662;
 XX

DT 03-APR-1996 (first entry)

DE Fusion protein contg. proinsulin sequence 3.

KW Proinsulin; post-translational modification; recombinant production;
 KW protein folding; conformation.

OS Synthetic.

XX Key Location/Qualifiers
 XX Region 41..44
 FT

FT	/label= R2
FT	/note= "a peptide of 4 amino acids"
FT	45..74
FT	/label= R1-(B2-B29)-Y
FT	/note= "human insulin B-chain"
FT	75
FT	/label= X
FT	76..96
FT	/label= GLY-(A2-A20)-R3
FT	/note= "human insulin A-chain"
XX	
XX	EP668292-A2.
PD	23-AUG-1995.
XX	
PE	09-FEB-1995; 95EP-00101748.
XX	
XX	18-FEB-1994; 94DE-04405179.
PA	(PARH) ROECHST AG.
PI	Obermeier R, Gerl M, Ludwig J, Sabel W;
DR	WPI, 1995-284754/38.
XX	
PT	Isolation of insulin that is correctly post-translationally processed -
PT	by reacting proinsulin with a mercaptan in the presence of a chaotropic
PT	agent and purification after absorption to hydrophobic resin.
XX	
PS	Example 2; Page 8; 16pp; German.
XX	
CC	The present sequence is that of a fusion protein, produced in E.coli
CC	which contains an example of a proinsulin molecule corresp. to the
CC	general formula R2-R1-(B2-B29)-Y-X-Gly-(A2-A20)-R3 (II). In formula (II),
CC	X = Lys, Arg or a peptide of 2-35 amino acids contg. Lys or Arg at the N-
CC	and C-termini; Y = a natural amino acid; R1 = Phe or a bond; R2 = H, Arg,
CC	Lys, a peptide of 2-45 amino acids contg. Arg or Lys at the N- and C-
CC	termini; R3 = a natural amino acid; (A2-A20) and (B2-B29) are the insulin
CC	A- and B-chain sequences from human or other insulin. The proinsulin
CC	molecule, released by cyanogen bromide, is reacted with mercaptan at a
CC	ratio of 2-10 SH residues of mercaptan per Cys residue of proinsulin. The
CC	reaction takes place in the presence of a chaotropic auxiliary agent at
CC	pH 10-11 and results in proinsulin with correctly linked cystine bridges.
CC	Reaction with trypsin and opt. carboxypeptidase B yields correctly folded
CC	insulin. The insulin is isolated by absorption on a hydrophobic resin
XX	
SO	Sequence 96 AA:
Query March	100.0%; Score 294; DB 2; Length 96;
Best Local Similarity	100.0%; Pred. No.2.5e-26;
Matches 52; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVECCCTSICTSLVQLENYCN 52
DB	45 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVECCCTSICTSLVQLENYCN 96
RESULT 7	
ID	AA42860 standard; protein; 107 AA.
AC	AA42860;
DT	19-JAN-2000 (first entry)
XX	
DE	hGH-mini-proinsulin chimeric protein.
KM	Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
OS	conformation; chimeric protein; cleavable; recombinant; production;
OS	yield.
OS	Synthetic.
OS	Hom sapiens.

XX	FN		MO9950302-AI.	
XX	PD	07-OCT-1999.		
XX	PB			
XX	PF	31-MAR-1998;	98WO-CN000052.	
XX	PR	31-MAR-1998;	98WO-CN000052.	
XX	PA	(TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.		
PI	P1	Gan Z;		
XX	DD	WPI; 1999-610839/52.		
PT	PT	New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.		
XX	PS	Claim 13; Page 30; 46pp; English.		
XX	ES	This sequence represents a chimeric protein, hGH-mini-Proinsulin. This chimeric protein contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AA42855, a cleavable peptide linker (AAV42857), and a human insulin precursor comprising insulin A and B chains (AAV42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary. The IMC sequences not only resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphydryls and related purification steps can thus be eliminated, along with the use of high concentrations of mercapтан or the use of hydrophobic absorbent resins		
CC	SQ	Sequence 107 AA;		
CC	Query Match	100.0%; Score 294; DB 2; Length 107;		
CC	Best Local Similarity	100.0%; Pred. No. 2, 8e-26;		
CC	Matches	52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
CY	1 FVNQHICGSHLVEALYLVCGERGFFYTPKTRIGIVQECCISGLYLENYCN 52			
DB	56 FVNQHICGSHLVEALYLVCGERGFFYTPKTRIGIVQECCISGLYLENYCN 107			
RESULT 8	ID	AAR98897 standard; protein; 116 AA.		
XX	AC	AAR98897;		
XX	DT	03-FEB-1997 (first entry)		
XX	DE	SOD-proinsulin hybrid polypeptide.		
XX	KM	Insulin; proinsulin; hybrid polypeptide; protein folding; enzymatic cleavage; cyanogen bromide; sulphhydryls.		
XX	KX	Homo sapiens.		
XX	PN	WC9620724-AI.		
XX	PD	11-JUL-1996.		
XX	PF	29-DEC-1994; 94WC-US013268.		

```

XX 29-DEC-1994; 94WO-US013268.
XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX
XX Hartman JR, Mendelovitz S, Gorecki M;
XX WPI; 1996-333766/33.
XX N-PSDB; AAQ34670.
XX
XX Recombinant insulin prodn. by correctly folding pro-insulin hybrid
XX polypeptide - then enzymatic cleavage of folded product, does not require
XX sulphite protection of SH nor use of cyanogen bromide.
XX
XX Example 1B; Fig 7, 69pp; English.
XX
XX A new method for the production of recombinant human insulin comprises
XX folding a hybrid polypeptide comprising proinsulin under conditions that
XX permit correct disulphide bond formation and subjecting that folded
XX protein to enzymatic cleavage. The insulin produced can then be purified.
XX This sequence is a SOD-insulin B chain-Arg-insulin A chain hybrid
XX polypeptide and is encoded by the plasmid construct pDBAST-LAR.
XX Transformation of the proper E.coli host cells with pDBAST-LAR results in
XX the efficient expression of the proinsulin hybrid polypeptide, useful for
XX human insulin production. The method produces recombinant human insulin
XX identical to the natural hormone. Hazardous and cumbersome procedures
XX involving cyanogen bromide and sulphiteolysis to protect SH groups are
XX avoided since the entire hybrid polypeptide folds efficiently to the
XX native structure even with the leader attached and Cys unprotected
XX
XX Sequence 116 AA;
SQ
Query Match 100.0%; Score 294; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.1e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 FVNHGCGSHLVKALYVCGERGFFYPTKRGIVGCCSTICSLVQLENYCN 52
65 FVNHGCGSHLVKALYVCGERGFFYPTKRGIVGCCSTICSLVQLENYCN 116
RESULT 9
AA71692 standard; protein; 137 AA.
XX
XX AAR71692;
XX
XX 25-MAR-2003 (revised)
XX DT 20-NOV-1995 (first entry)
XX
XX Mating factor alpha 1-Insulin precursor ArgB31.
XX
XX Human insulin precursor ArgB31; diabetes; zinc ion complex;
XX mating factor alpha 1.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT 1..85
XX FT /label= mating factor alpha-1
XX FT Peptide 86..116
XX FT /label= B-chain
XX FT Peptide 117..137
XX FT /label= A-chain
XX
XX WO9507931-A1.
XX
XX 23-MAR-1995.
XX
XX 16-SEP-1994; 94WO-DK000347.
XX
XX 17-SEP-1993; 93DK-00001044.
XX PR 02-FEB-1994; 94US-00190829.

```

```

XX (NOVO ) NOVO-NORDISK AS.
XX
XX Havelund S, Halstrom JB, Jonassen I, Andersen AS, Markussen J;
XX WPI; 1995-131314/17.
XX N-PSDB; AAQ86425.
XX
XX Acylated insulin deriv. which may be present as a zinc ion complex - is
XX used to treat diabetes and is rapid acting.
XX
XX Example 5; Page 78; 100pp; English.
XX
XX AAQ86425 encodes AAR71692 mating factor alpha 1-Insulin precursor ArgB31.
XX ArgB31 comprises the B and A chains of a claimed human insulin
XX derivative. In the final claimed compen. they are covalently connected
XX via disulphide bonds between Cys residues A7/B7 and A20/B19. The
XX derivative, which may be present as a zinc ion complex, can be used as a
XX fast action treatment for diabetes. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
XX Sequence 137 AA;
SQ
Query Match 100.0%; Score 294; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.7e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 FVNHGCGSHLVKALYVCGERGFFYPTKRGIVGCCSTICSLVQLENYCN 52
66 FVNHGCGSHLVKALYVCGERGFFYPTKRGIVGCCSTICSLVQLENYCN 137
RESULT 10
AAR71694 standard; protein; 145 AA.
XX
XX AAR71694;
XX
XX 25-MAR-2003 (revised)
XX DT 20-NOV-1995 (first entry)
XX
XX Mating factor alpha 1-Insulin precursor ArgB1, ArgB31 N-terminal.
XX
XX Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex;
XX mating factor alpha 1; N-terminal EEAFAEAR.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT 1..85
XX FT /label= mating factor alpha-1
XX FT Peptide 86..93
XX FT /label= N-terminal peptide
XX FT Peptide 94..124
XX FT /label= B-chain
XX FT Peptide 125..145
XX FT /label= A-chain
XX
XX WO9507931-A1.
XX
XX 23-MAR-1995.
XX
XX 16-SEP-1994; 94WO-DK000347.
XX
XX 17-SEP-1993; 93DK-00001044.
XX PR 02-FEB-1994; 94US-00190829.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Havelund S, Halstrom JB, Jonassen I, Andersen AS, Markussen J;
XX WPI; 1995-131314/17.
XX N-PSDB; AAQ86425.

```

XX Acylated insulin deriv. which may be present as a zinc ion complex - is
 PT used to treat diabetes and is rapid acting.
 XX
 PS Example 5; Page 82-83; 100pp; English.
 CC AAQ86432 encodes AAR71695 mating factor alpha 1-insulin precursor ArgB1,
 CC ArgB31 N-terminal EBAEAAR. The insulin precursor comprises the B and A
 CC chains of a claimed human insulin derivative preceded by the N-terminal
 CC amino acids EBAEAAR. In the final claimed compsn. they are covalently
 CC connected via disulphide bonds between Cys residues A7/B7 and A20/B19.
 CC The derivative, which may be present as a zinc ion complex, can be used
 CC as a fast action treatment for diabetes. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 145 AA;
 Query Match 100.0%; Score 294; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 3.9e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 FVNOHCGSHVEALYVCGRGFFPKTRGIVECCSTICSLYOLENYCN 52
 Db 94 FVNOHCGSHVEALYVCGRGFFPKTRGIVECCSTICSLYOLENYCN 145
 RESULT 11
 AAR71695
 ID AAR71695 standard; protein; 146 AA.
 XX
 AC AAR71695;
 XX
 DT 25-MAR-2003 (revised)
 DT 20-NOV-1995 (first entry)
 XX
 DE Mating factor alpha 1-Insulin precursor ArgB1, ArgB31 N-terminal.
 XX
 KM Human insulin precursor ArgB1, ArgB31; diabetes; Zinc ion complex;
 KM mating factor alpha 1; N-terminal EBAEAAR.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..85
 FT Peptide /label= mating factor alpha-1
 FT Peptide 86..94
 FT Peptide /label= N-terminal peptide
 FT Peptide 95..125
 FT Peptide /label= B-chain
 FT Peptide 126..146
 FT Peptide /label= A-chain
 XX
 PN WO9507931-A1.
 XX
 PD 23-MAR-1995.
 XX
 PF 16-SEP-1994; 94WO-DK000347.
 XX
 PR 17-SEP-1993; 93DK-00001044.
 PR 02-FEB-1994; 94US-00190829.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Havelund S, Halstrom JB, Jonassen I, Andersen AS, Markussen J;
 XX
 DR WPI; 1995-131314/17.
 DR N-PSDB; AAQ86432.
 XX
 PT Acylated insulin deriv. which may be present as a zinc ion complex - is
 PT used to treat diabetes and is rapid acting.
 XX
 PS Example 6; Page 85; 100pp; English.

CC AAQ86432 encodes AAR71695 mating factor alpha 1-insulin precursor ArgB1,
 CC ArgB31 N-terminal EBAEAAR. The insulin precursor comprises the B and A
 CC chains of a claimed human insulin derivative preceded by the N-terminal
 CC amino acids EBAEAAR. In the final claimed compsn. they are covalently
 CC connected via disulphide bonds between Cys residues A7/B7 and A20/B19.
 CC The derivative, which may be present as a zinc ion complex, can be used
 CC as a fast action treatment for diabetes. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 146 AA;
 Query Match 100.0%; Score 294; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 3.9e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 FVNOHCGSHVEALYVCGRGFFPKTRGIVECCSTICSLYOLENYCN 52
 Db 95 FVNOHCGSHVEALYVCGRGFFPKTRGIVECCSTICSLYOLENYCN 146
 RESULT 12
 AA42861
 ID AA42861 standard; protein; 150 AA.
 XX
 AC AA42861;
 XX
 DT 19-JAN-2000 (first entry)
 DT
 DE Chimeric protein, SEQ ID 7.
 XX
 KM Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
 KM conformation; chimeric protein; cleavable; recombinant; production;
 KM yield.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9503032-A1.
 PD 07-OCT-1999.
 XX
 PF 31-MAR-1998; 98WO-CN000052.
 PR 31-MAR-1998; 98WO-CN000052.
 XX
 (TONG-) TONGHUA GANTBCH BIOTECHNOLOGY LTD.
 XX
 Gan Z;
 DR WPI; 1999-610839/52.
 XX
 PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.
 XX
 PS Claim 14; Page 30-31; 46pp; English.
 XX
 CC This sequence represents a chimeric protein, which contains an N-terminal
 CC fragment of human growth hormone (hGH) of the sequence given in AA42861,
 CC a cleavable peptide linker (AA42857), and a human insulin precursor
 CC comprising insulin A and B chains (AA42859). The hGH portion of the
 CC chimeric protein acts as an intramolecular chaperone (IMC) for the
 CC insulin precursor, enabling it to fold correctly. The cleavable peptide
 CC linker has a C-terminal Arg residue which enables the hGH portion of the
 CC chimeric protein to be removed after folding has taken place. Production
 CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
 CC provide human insulin with correctly linked cysteine bridges with fewer
 CC necessary procedural steps, and hence resulting in a higher yield of
 CC human insulin. The IMC sequences not only protect insulin sequences from
 CC intracellular degradation by a microorganism host, but also promote the
 CC folding of the fused insulin precursor, facilitate the solubility of the
 CC fusion protein and decrease the intermolecular interactions among the
 CC fusion proteins, thus allowing folding of the fused insulin precursor at
 CC commercially useful high concentrations. The procedural steps of cyanogen

CC bromide cleavage, oxidative sulphytolysis and related purification steps
CC can thus be eliminated, along with the use of high concentrations of
CC mercaptan or the use of hydrophobic absorbent resins

XX Sequence 150 AA;

Query Match 100.0%; Score 294; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 4e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYVCGERGFYPTKTRGIVEOCCTSGSLYLENYCN 52
DB 99 FVNOHLCGSHLVEALYVCGERGFYPTKTRGIVEOCCTSGSLYLENYCN 150

RESULT 13

AA04582
ID AA04582 standard; protein; 57 AA.

AC AA04582;

DT 09-SEP-2004 (revised)

DT 25-MAR-2003 (revised)

DT 14-SEP-1990 (first entry)

DE Proinsulin analogue with a Lys residue linking the A and B chains.

KW insulin fusion protein; pro-insulin analogue; tendamistate;

OS Lys-Lys bridge; ds.

XX Synthetic.

PH Key Location/Qualifiers

FT Peptide 1..35 "Insulin B chain"

FT MISC-difference 36 /note= "Lys residue linking insulin B chain to A chain"

FT Peptide 37..57 /note= "Insulin A chain"

XX EP367163-A.

XX PD 09-MAY-1990.

PF 28-OCT-1989; 89EP-00120056.

PR 03-NOV-1988; 88DE-03837273.

PR 19-AUG-1989; 89DE-03927449.

XX (PARH) HOECHST AG.

PI Koller KP, Riess GJ, Uhlmann E, Wallmeier H;

DR WPI; 1990-141149/19.

DR N-PSDB; AA004335.

XX New insulin fusion proteins - comprise pro-insulin analogue linked to

XX tendamistate.

PS Disclosure; Page 5; 8pp; German.

XX This sequence is joined to the C-terminus of an N-terminal fragment

XX comprising opt. modified tendamistate. This fusion protein may be

XX converted into human insulin using known methods. The synthetic gene was

XX prepared by the phosphoramidite method. See also AA004336. (Updated on 25

XX -MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI

XX field.)

XX Revised record issued on 09-SEP-2004 : Correction to pages and features

XX Sequence 57 AA;

XX Query Match 99.0%; Score 291; DB 2; Length 57;

Best Local Similarity 98.1%; Pred. No. 3.3e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYVCGERGFYPTKTRGIVEOCCTSGSLYLENYCN 52
DB 6 FVNOHLCGSHLVEALYVCGERGFYPTKTRGIVEOCCTSGSLYLENYCN 57

RESULT 14

AA01899
ID AA01899 standard; protein; 52 AA.

AC AA01899;

DT 25-MAR-2003 (revised)

DT 22-JUL-1991 (first entry)

DE Example of human insulin precursor.

KW Human insulin; diabetes; transpeptidation.

OS Homo sapiens.

PN EP427296-A.

PD 15-MAY-1991.

PF 29-MAY-1985; 90EP-00121887.

PR 30-MAY-1984; 84DK-00002665.

PR 08-FEB-1985; 85DK-00000582.

XX (NOVO) NOVO-NORDISK AS.

PI Markussen J, Fill N, Ammerer G, Hansen MT, Thim L, Norris K;

PI Volgt HO;

DR WPI; 1991-141828/20.

XX Human insulin precursors - expressed with correctly positioned

XX di-sulphide bridges giving improved resistance to proteolysis.

XX Claim 3; Page 18; 28pp; English.

XX This human insulin precursor has correctly positioned disulphide bridges

XX between the A and B chains and is more resistant to proteolytic digestion

XX than prior art insulin precursors. Yeast strains transformed with DNA

XX encoding this precursor can be cultured to secrete it in high yields. The

XX precursor can be converted into mature human insulin by transpeptidation.

XX See also AA01897-98. (Updated on 25-MAR-2003 to correct PF field.)

XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 52 AA;

Query Match 97.6%; Score 287; DB 2; Length 52;

Best Local Similarity 96.2%; Pred. No. 8.8e-26;

Matches 50; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYVCGERGFYPTKTRGIVEOCCTSGSLYLENYCN 52

DB 1 FVNOHLCGSHLVEALYVCGERGFYPTKTRGIVEOCCTSGSLYLENYCN 52

RESULT 15

AA05883

ID AA05883 standard; protein; 53 AA.

AC AA05883;

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 26-JUN-1995 (first entry)

Search completed: November 2, 2004, 20:11:47
Job time : 51.1771 secs

Sat Nov 6 18:59:27 2004

us-10-054-873-5.rai

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: November 2, 2004, 20:02:41 / Search time 12.6642 Seconds
(Without alignments)
272.306 Million cell updates/sec

Title: US-10-054-873-5

Perfect score: 294

Sequence: 1 FVNQHLGSHLVEALYLVCG.....IVEQCTSLICSLYLENYCN 52

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	100.0	56	1	US-08-160-376A-7
2	294	100.0	56	1	US-08-389-487-11
3	294	100.0	63	1	US-08-160-376A-6
4	294	100.0	66	1	US-08-291-0608-5
5	294	100.0	96	1	US-08-160-376A-5
6	294	100.0	96	1	US-08-389-487-8
7	294	100.0	137	1	US-08-400-256-39
8	294	100.0	137	3	US-08-975-365-39
9	294	100.0	145	1	US-08-400-256-45
10	294	100.0	145	3	US-08-975-365-45
11	294	100.0	146	1	US-08-400-256-48
12	294	100.0	146	3	US-08-975-365-48
13	291	99.0	57	1	US-08-030-731A-44
14	283.5	96.4	53	1	US-08-233-617-4
15	283.5	96.4	53	3	US-08-981-988A-42
16	278.5	94.7	51	4	US-09-477-924-3
17	278.5	94.7	51	4	US-09-723-981-3
18	278.5	94.7	51	4	US-09-723-981-3
19	277.5	94.4	53	1	US-08-233-617-3
20	277.5	94.2	55	3	US-08-900-574-3
21	276.5	94.0	66	3	US-08-900-574-6
22	276.5	94.0	66	3	US-08-900-574-5
23	276.5	94.0	67	3	US-08-981-988A-1
24	276.5	94.0	67	3	US-08-981-988A-5
25	275.5	93.9	67	3	US-08-900-574-7
26	275.5	93.7	53	3	US-09-261-853-2
27	273.5	93.7	63	1	US-08-468-674B-71

28	275.5	93.7	65	1	US-08-780-571-71	Sequence 71, Appl
29	275.5	93.7	89	1	US-08-468-674B-41	Sequence 41, Appl
30	275.5	93.7	89	1	US-08-780-571-41	Sequence 41, Appl
31	275.5	93.7	91	1	US-08-468-674B-45	Sequence 45, Appl
32	275.5	93.7	91	1	US-08-780-571-45	Sequence 45, Appl
33	275.5	93.7	104	1	US-08-400-256-15	Sequence 15, Appl
34	275.5	93.7	104	3	US-08-975-365-15	Sequence 15, Appl
35	275.5	93.7	117	3	US-09-012-669F-37	Sequence 37, Appl
36	275.5	93.7	124	1	US-08-446-646-3	Sequence 3, Appl
37	275.5	93.7	124	1	US-09-012-669F-36	Sequence 3, Appl
38	275.5	93.7	124	4	US-09-894-711-18	Sequence 18, Appl
39	275.5	93.7	138	4	US-08-932-082-19	Sequence 19, Appl
40	275.5	93.7	138	4	US-09-861-687-19	Sequence 19, Appl
41	275.5	93.7	140	1	US-08-400-256-33	Sequence 33, Appl
42	275.5	93.7	140	1	US-08-400-256-33	Sequence 42, Appl
43	275.5	93.7	140	3	US-08-975-365-33	Sequence 33, Appl
44	275.5	93.7	140	3	US-08-975-365-42	Sequence 42, Appl
45	273.5	93.0	67	3	US-08-981-988A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-160-376A-7
Sequence 7, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeyer, Ranier
APPLICANT: Geil, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Geroni, Esq.
STREET: Rt. 202-206 No. 5473049ct/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOE 92/P 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-2255
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-7
Query Match 100.0%; Score 294; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 9.4e-29;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNOHLCGSHLVEALYVCGERGFFYPTKTRGIVEOCCCTISICSLYOLENYCN 52
DB 5 FVNOHLCGSHLVEALYVCGERGFFYPTKTRGIVEOCCCTISICSLYOLENYCN 56

RESULT 2

US-08-389-487-11
Sequence 11, Application US/08389487

Patent No. 5663291

GENERAL INFORMATION:

APPLICANT: Obermeyer, Rainer

APPLICANT: Gerl, Martin

APPLICANT: Ludwig, Jürgen

APPLICANT: Sabel, Walter

TITLE OF INVENTION: Process for Obtaining Insulin Having

TITLE OF INVENTION: Correctly Linked Cysteine Bridges

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS: Correctly Linked Cysteine Bridges

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: United States of America

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/389,487

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Einaudi, Carol P.

REGISTRATION NUMBER: 32,220

REFERENCE/DOCKET NUMBER: 02481.1424-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 56 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-389-487-11

Query Match 100.0%; Score 294; DB 1; Length 56;

Best Local Similarity 100.0%; Pred. No. 9.4e-29;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNOHLCGSHLVEALYVCGERGFFYPTKTRGIVEOCCCTISICSLYOLENYCN 52
DB 5 FVNOHLCGSHLVEALYVCGERGFFYPTKTRGIVEOCCCTISICSLYOLENYCN 56

RESULT 3

US-08-160-376A-6
Sequence 6, Application US/08160376A

Patent No. 5473049

GENERAL INFORMATION:

APPLICANT: Obermeyer, Rainer

APPLICANT: Gerl, Martin

APPLICANT: Ludwig, Jürgen

APPLICANT: Sabel, Walter

TITLE OF INVENTION: Process for Obtaining Proinsulin

TITLE OF INVENTION: Possessing Correctly Linked

TITLE OF INVENTION: Cysteine Bridges

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049CN/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

COMPUTER: IBM 386

OPERATING SYSTEM: WINDOWS 3.1

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/160,376A

FILING DATE: December 1, 1993

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GE P 4240420.7

FILING DATE: December 2, 1992

ATTORNEY/AGENT INFORMATION:

NAME: Barbara V. Maurer, Esq.

REGISTRATION NUMBER: 31,287

REFERENCE/DOCKET NUMBER: HOE 92/P 384

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 231-4079

TELEFAX: (908) 231-2255

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 63 Amino Acids

TYPE: Amino Acid (AA)

TOPOLOGY: not relevant

US-08-160-376A-6

Query Match 100.0%; Score 294; DB 1; Length 63;

Best Local Similarity 100.0%; Pred. No. 1.1e-28;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNOHLCGSHLVEALYVCGERGFFYPTKTRGIVEOCCCTISICSLYOLENYCN 52
DB 12 FVNOHLCGSHLVEALYVCGERGFFYPTKTRGIVEOCCCTISICSLYOLENYCN 63

RESULT 4

US-08-291-060B-5
Sequence 5, Application US/08291060B

Patent No. 5728543

GENERAL INFORMATION:

APPLICANT: Dorschug, Michael

APPLICANT: Koller, Klaus-Peter

APPLICANT: Marguardt, Rudiger

APPLICANT: Meibes, Johannes

TITLE OF INVENTION: An Enzymatic Process for the

TITLE OF INVENTION: Conversion of Preproinsulins Into Insulins

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/291,060B

FILING DATE: 08-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Einaudi, Carol P.

REGISTRATION NUMBER: 32,220

Sat Nov 6 18:59:27 2004

us-10-054-873-5.rai

Page 3

REFERENCE/DOCKET NUMBER: 02481.1105-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4366
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-291-0608-5

Query Match 100.0%; Score 294; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTISCSLYOLENYCN 52
Db 15 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTISCSLYOLENYCN 66

RESULT 5
US-08-160-376A-5
Sequence 5, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeyer, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
CYSTINE BRIDGES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Geroni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOE 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-5

Query Match 100.0%; Score 294; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.7e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTISCSLYOLENYCN 52

Db 45 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTISCSLYOLENYCN 96

RESULT 6
US-08-389-487-8
Sequence 8, Application US/08389487
Patent No. 5663291
GENERAL INFORMATION:
APPLICANT: Obermeyer, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Insulin Having
TITLE OF INVENTION: Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fannegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunnet
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/389,487
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Binardi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481.1424-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-389-487-8

Query Match 100.0%; Score 294; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.7e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTISCSLYOLENYCN 52
Db 45 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTISCSLYOLENYCN 96

RESULT 7
US-08-400-256-39
Sequence 39, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497 No. 5750497/disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor

CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-400-256-39

Query Match 100.0%; Score 294; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 2,4e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FVNHLCGSHLYEALYVCGERGFFYTPKRGIVGECCTCSISLYOLENYCN 52
86 FVNHLCGSHLYEALYVCGERGFFYTPKRGIVGECCTCSISLYOLENYCN 137

RESULT 8
US-08-975-365-39
Sequence 39, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6011007 No. 6011007disk of No. 6011007ch America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,365
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-975-365-39

Query Match 100.0%; Score 294; DB 3; Length 137;
Best Local Similarity 100.0%; Pred. No. 2,4e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FVNHLCGSHLYEALYVCGERGFFYTPKRGIVGECCTCSISLYOLENYCN 52
86 FVNHLCGSHLYEALYVCGERGFFYTPKRGIVGECCTCSISLYOLENYCN 137

RESULT 9
US-08-400-256-45
Sequence 45, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5750497 No. 5750497disk of No. 5750497ch America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-400-256-45

Query Match 100.0%; Score 294; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 2,6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FVNHLCGSHLYEALYVCGERGFFYTPKRGIVGECCTCSISLYOLENYCN 52
94 FVNHLCGSHLYEALYVCGERGFFYTPKRGIVGECCTCSISLYOLENYCN 145

RESULT 10
US-08-975-365-45

Sat Nov 6 18:59:27 2004

us-10-054-873-5.rai

Page 5

```
Sequence 45, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,365
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-975-365-45

Query Match 100.0%; Score 294; DB 3; Length 145;
Best local Similarity 100.0%; Pred. No. 2.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALIVCGERGFYTPKTRIGIVECCCTSLVQLNENCN 52
DB 94 FVNHLCGSHLVEALIVCGERGFYTPKTRIGIVECCCTSLVQLNENCN 145

RESULT 11
US-08-400-256-48
Sequence 48, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: No. 57504970 No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
```

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-400-256-48

Query Match 100.0%; Score 294; DB 1; Length 146;
Best local Similarity 100.0%; Pred. No. 2.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALIVCGERGFYTPKTRIGIVECCCTSLVQLNENCN 52
DB 95 FVNHLCGSHLVEALIVCGERGFYTPKTRIGIVECCCTSLVQLNENCN 146

RESULT 12
US-08-975-365-48
Sequence 48, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,365
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
```

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-975-365-48

Query Match 100.0%; Score 294; DB 3; Length 146;
Best Local Similarity 100.0%; Pred. No. 2,6e-28; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLYEALYIVCGRGFFYPKTRGIVECCCTISCSLYOLENYCN 52
DB 95 FVNHLCGSHLYEALYIVCGRGFFYPKTRGIVECCCTISCSLYOLENYCN 146

RESULT 13
US-08-030-731A-44
Sequence 44, Application US/08030731A

GENERAL INFORMATION:

APPLICANT: Koller, Klaus-Peter
APPLICANT: Riess, Guenther Johannes
APPLICANT: Uhlmann, Eugen
APPLICANT: Wallemier, Holger
TITLE OF INVENTION: Processes for the Preparation of Foreign
TITLE OF INVENTION: Proteins in Streptomyces
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,731A
FILING DATE: 12-MAR-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/189,840
FILING DATE: 03-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/430,622
FILING DATE: 01-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/687,610
FILING DATE: 19-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,757
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 37 14 866.4
FILING DATE: 05-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 38 37 273.8
FILING DATE: 03-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 39 27 449.7
FILING DATE: 19-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 40 12 818.0
FILING DATE: 21-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Kirschner Michael K.
REGISTRATION NUMBER: 34,851
REFERENCE/DOCKET NUMBER: 02481-0593-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-030-731A-44

Query Match 99.0%; Score 291; DB 1; Length 57;
Best Local Similarity 98.1%; Pred. No. 1,7e-27;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLYEALYIVCGRGFFYPKTRGIVECCCTISCSLYOLENYCN 52
DB 6 FVNHLCGSHLYEALYIVCGRGFFYPKTRGIVECCCTISCSLYOLENYCN 57

RESULT 14
US-08-233-617-4
Sequence 4, Application US/08233617

GENERAL INFORMATION:

APPLICANT: Obermeier, Rainer
APPLICANT: Sabel, Walter
APPLICANT: Dell, Peter
APPLICANT: Gelsen, Karl
TITLE OF INVENTION: Amphiphous Monospherical Forms of Insulin
TITLE OF INVENTION: Derivatives
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,617
FILING DATE: 25-APR-1994
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 43 13 702.4
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Carol P. Einaudi
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481.1374-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
US-08-233-617-4

Query Match 96.4%; Score 283.5; DB 1; Length 53;
Best Local Similarity 98.1%; Pred. No. 1,7e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNHLCGSHLYEALYIVCGRGFFYPKTRGIVECCCTISCSLYOLENYCN 52

Sat Nov 6 18:59:27 2004

us-10-054-873-5.rat

Page 7

Db 1 FVNQHCGSHLVKALYVCGERGFYTPKTRGIVECCCTISCSLYOLENYCN 53

RESULT 15
US-08-981-988A-42
; Sequence 42, Application US/08981988A
; Patent No. 6337194
; GENERAL INFORMATION:
; APPLICANT: Vittal Maliya Scientific Research Foundation
; APPLICANT: The University of Leicester
; TITLE OF INVENTION: Insulin
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VITTAL MALIYA SCIENTIFIC RESEARCH FOUNDATION
; STREET: K. R. ROAD
; CITY: BANGALORE
; COUNTRY: INDIA
; ZIP: 560 004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,988A
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9513967.1
; FILING DATE: 08-JUL-1995
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-981-988A-42

Query Match 96.4%; Score 283.5; DB 3; Length 53;
Best Local Similarity 98.1%; Pred. No. 1.7e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 FVNQHCGSHLVKALYVCGERGFYTPKTRGIVECCCTISCSLYOLENYCN 52
Db 1 FVNQHCGSHLVKALYVCGERGFYTPKTRGIVECCCTISCSLYOLENYCN 53

Search completed: November 2, 2004, 20:24:35
Job time: 13.6642 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:59:41 / Search time 9.594 seconds

(without alignments)
521.495 Million cell updates/sec

Title: US-10-054-873-5

Perfect score: 294

Sequence: 1 FVNHGSHLVKALYVCG.....IVEQCTGVSICLYOLENYCN 52

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	273.5	93.0	51	1 INEL	insulin - elephant
2	273.5	93.0	51	1 INWH	insulin - finback
3	273.5	93.0	51	1 INWH	insulin - sperm wh
4	273.5	93.0	96	2 PC7082	epidermal growth f
5	271.5	92.3	51	1 INWH	insulin - hamster
6	268.5	91.3	51	1 INWSP	insulin - Egyptian
7	267.5	91.0	51	2 A59151	insulin precursor
8	267.5	90.8	110	1 IPHU	insulin precursor
9	267.5	90.8	110	2 B42179	insulin precursor
10	267.5	90.8	110	2 U00178	insulin precursor
11	267.5	90.8	110	2 A42179	insulin precursor
12	263.5	89.6	51	1 INWMA	insulin - Arabian
13	263.5	89.6	51	1 INGT	insulin - goat
14	263.5	89.6	51	1 INWHS	insulin - sei whale
15	263.5	89.6	84	1 IPPG	insulin precursor
16	263.5	89.3	110	1 INRB	insulin precursor
17	262.5	89.3	51	1 INCT	insulin precursor
18	262.5	89.1	110	1 IPDG	insulin precursor
19	261.5	88.9	51	1 INWKO	insulin - common s
20	260.5	88.4	110	2 A48166	insulin precursor
21	258.5	87.9	105	1 IPBO	insulin precursor
22	256.5	87.2	51	2 J00362	insulin - North Am
23	252.5	85.9	77	1 INSH	insulin precursor
24	252.5	85.7	86	1 IPHO	insulin precursor
25	251.5	85.5	51	1 INCB	insulin - Chichil
26	251.5	85.4	108	2 A39883	insulin precursor
27	250.5	85.0	108	2 INMS1	insulin precursor
28	249.5	84.7	110	1 IPRT1	insulin - goose
29	248.5	84.5	51	1 INGS	insulin - goose

30	248	84.4	110	1 INMS2	insulin 2 precursor
31	248	84.4	110	1 IPRT2	insulin 2 precursor
32	246	83.7	52	2 S44470	insulin 12 - North
33	246	83.7	52	2 S44469	insulin 11 - North
34	244.5	83.2	51	1 INOS	insulin - ostrich
35	244.5	83.2	51	1 INTK	insulin - turkey
36	244.5	83.2	51	1 A61129	insulin - black-be
37	244.5	83.2	51	1 INPQ	insulin - crested
38	244.5	83.2	51	2 A60614	insulin - slider t
39	236.5	81.1	103	2 IS1221	insulin precursor
40	238	81.0	52	2 S61361	insulin - Amphibia
41	235.5	80.1	51	2 S63590	insulin - duckbill
42	234.5	79.8	107	1 IPCH	insulin precursor
43	233.5	79.4	81	1 IPDK	insulin precursor
44	231.5	78.7	51	1 INAQ	insulin - American
45	231	78.6	52	1 INGA	insulin - alligato

ALIGNMENTS

RESULT 1

INEL

C:Species: Elephantidae Gen. sp. (elephant)

C:Date: 24-Apr-1984 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999

C:Accession: A01584

R:Smith, L.F.

Am. J. Med. 40, 662-666, 1966

A:Title: Species variation in the amino acid sequence of insulin.

A:Reference number: A90029; PMID:66160119; PMID:5949593

A:Accession: A01584

A:Residues: 1-30;31-51 <SWT>

A:Note: the species of elephant is not given, but it is most probably the Indian eleph

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-30/Domain: Insulin chain B #status experimental <BCH>

F:31-51/Product: Insulin #status experimental <MAT>

F:31-51/Domain: Insulin chain A #status experimental <ACH>

F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 93.0%; Score 273.5; DB 1; Length 51;

Best Local Similarity 94.2%; Pred. No. 1.5e-24; Matches 49; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 FVNHGSHLVKALYVCGERGFYPTKRGIVEQCTGVSICLYOLENYCN 52
DB 1 FVNHGSHLVKALYVCGERGFYPTKRGIVEQCTGVSICLYOLENYCN 51

RESULT 2

INWH

C:Species: Baleenoptera physalus (finback whale, common forqual)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C:Accession: A91918

R:Hama, H.; Tiltan, K.; Sakaki, S.; Narita, K.

J. Biochem. 56, 285-293, 1964

A:Title: The amino acid sequence in fin-whale insulin.

A:Reference number: A91918

A:Accession: A91918

A:Molecule type: protein

A:Residues: 1-30;31-51 <HAM>

A:Cross-references: UNIPROT:P01312

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-30/Domain: Insulin chain B #status experimental <BCH>

F:31-51/Product: Insulin #status experimental <MAT>

F:31-51/Domain: Insulin chain A #status experimental <ACH>

F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 93.0%; Score 273.5; DB 1; Length 51;

Best Local Similarity 96.2%; Pred. No. 1.5e-24;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCSTISGLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKA-GIVEQCCSTISGLYLENYCN 51

RESULT 3

INMHP

insulin - sperm whale

C/Species: Physeta catodon (sperm whale)

C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C/Accession: A93142; A90082

R/Isihara, Y.; Saito, T.; Ito, Y.; Fujino, M.

A/Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pepsin.

A/Reference number: A93142

A/Accession: A93142

A/Molecule type: protein

A/Residues: 1-30,31-51 <ASH>

A/Cross-references: UNIPROT:P01312

R/Harris, J.I.; Sanger, F.; Naughton, M.A.

A/Title: Species differences in insulin.

A/Reference number: A90082

A/Accession: A90082

A/Molecule type: protein

A/Residues: 1-30,31-51 <HAR>

C/Superfamily: insulin

C/Keywords: hormone; pancreas

F/1-30/Domain: insulin chain B #status experimental <BCH>

F/1-30,31-51/Product: insulin #status experimental <MAT>

F/31-51/Domain: insulin chain A #status experimental <ACH>

F/7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 93.0%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 1.5e-24;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCSTISGLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKA-GIVEQCCSTISGLYLENYCN 51

RESULT 4

epidermal growth factor/single chain insulin fusion protein - Bacillus brevis (fragment)

C/Species: Bacillus brevis

C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004

C/Accession: PC7082; PC7083

R/Koh, M.; Hasegawa, H.; Ebisu, S.; Morihara, K.; Takagi, H.

A/Title: Use of Bacillus brevis for synthesis and secretion of Des-B30 single-chain human

A/Reference number: PC7082; PMID:2035834; PMID:10879487

A/Accession: PC7082

A/Molecule type: DNA

A/Residues: 1-96 <KOH>

A/Cross-references: UNIPROT:Q7M006

A/Accession: PC7083

A/Molecule type: protein

A/Residues: 19-28 <K02>

C/Genetics:

A/Gene: egf-scl

C/Superfamily: insulin

Query Match 92.9%; Score 273; DB 2; Length 96;
Best Local Similarity 96.2%; Pred. No. 3e-24;
Matches 50; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCSTISGLYLENYCN 52
DB 47 FVNQHLGSHLVEALYLVCGERGFFYTPK--GIVEQCCSTISGLYLENYCN 96

RESULT 5

INMHP

insulin - hamster

C/Species: Cricetinae gen. sp. (hamster)

C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999

C/Accession: A91456

R/Neelon, F.A.; Delcher, H.K.; Steinman, H.; Labovitz, H.E.

A/Title: Structure of hamster insulin: comparison with a tumor insulin.

A/Reference number: A91456

A/Accession: A91456

A/Molecule type: protein

A/Residues: 1-30,31-51 <NEE>

A/Cross-references: UNIPROT:Q7M0G1

C/Superfamily: insulin

C/Keywords: hormone; pancreas

F/1-30/Domain: insulin chain B #status experimental <BCH>

F/1-30,31-51/Product: insulin #status experimental <MAT>

F/31-51/Domain: insulin chain A #status experimental <ACH>

F/7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 92.3%; Score 271.5; DB 1; Length 51;
Best Local Similarity 94.2%; Pred. No. 2.6e-24;
Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCSTISGLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKS-GIVDQCCSTISGLYLENYCN 51

RESULT 6

INMSSP

insulin - Egyptian spiny mouse (tentative sequence)

C/Species: Acomys cahirinus (Egyptian spiny mouse)

C/Date: 13-Jul-1991 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004

C/Accession: A01591

R/Buenzli, H.F.; Humbel, R.E.

A/Title: Isolation and partial structural analysis of insulin from mouse (Mus musculus)

A/Reference number: A01591; PMID:72189454; PMID:5028210

A/Accession: A01591

A/Molecule type: protein

A/Residues: 1-30,31-51 <HUE>

A/Cross-references: UNIPROT:P01324

C/Superfamily: insulin

C/Keywords: hormone; pancreas

F/1-30/Domain: insulin chain B #status predicted <BCH>

F/1-30,31-51/Product: insulin #status predicted <MAT>

F/31-51/Domain: insulin chain A #status predicted <ACH>

F/7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 91.3%; Score 268.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 5.6e-24;
Matches 48; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCSTISGLYLENYCN 52
DB 1 FVBQHLGSHLVEALYLVCGERGFFYTPKS-GIVDQCCSTISGLYLENYCN 51

RESULT 7

A59151

insulin precursor - jack bean (fragments)

N/Alternate names: hypoglycemic agent; plant insulin

C/Species: Canavalia ensiformis (jack bean)

C/Date: 07-Dec-1999 #sequence_revision 07-Dec-1999 #text_change 10-Dec-1999

C/Accession: B59151; A59151

R/Oliveira, A.R.A.; Machado, O.L.T.; Gomes, V.M.; Xavier-Neto, J.; Pereira, A.C.P.; Vi

A/Title: Jack bean seed coat contains a protein with complete sequence homology to bo

A:Reference number: A59151
 A:Accession: B59151
 A:Molecule type: protein
 A:Residues: 1-30 <MACB>
 A:Cross-references: UNIPROT:Q7M217
 A:Accession: A59151
 A:Molecule type: protein
 A:Residues: 31-51 <MACA>
 C:Comment: The two chains are probably produced from the same precursor.
 C:Superfamily: Insulin
 F:1-30,31-51/Protein: Insulin #status experimental <MAT>
 F:1-30/Domain: chain B #status experimental <CHB>
 F:31-51/Domain: chain A #status experimental <CHA>
 F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 91.0%; Score 267.5; DB 2; Length 51;
 Best Local Similarity 92.3%; Pred. No. 7.3e-24;
 Matches 48; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 FVNQHLCGSHLVEAALPVCGEGFETPKTKGIVEQCCTISGLYOLENYCN 52
 DB 1 FVNQHLCGSHLVEAALPVCGEGFETPKTKA-GIVEQCCTISGLYOLENYCN 51

RESULT 8

Insulin precursor [validated] - human
 N:Alternate names: preproinsulin
 C:Species: Homo sapiens (man)
 C:Date: 23-Oct-1981 #sequence revision 23-Oct-1981 #text change 09-Jul-2004
 C:Accession: A93222; A94553; A99216; A94251; A93144; A92075; A91186; A01579; S58
 R:Bell, G.I.; Flicet, R.V.; Rutter, W.C.; Cordell, B.; Tischer, E.; Goodman, H.M.
 Nature 284, 26-32, 1980
 A:Title: Sequence of the human insulin gene.
 A:Reference number: A93222; MUID:80120725; PMID:6243748
 A:Accession: A93222
 A:Molecule type: DNA
 A:Residues: 1-110 <BRU>
 A:Cross-references: UNIPROT:P01308; GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R:Ulrich, A.; Dull, T.C.; Gray, A.; Brostius, J.; Sures, I.
 Science 209, 612-615, 1980
 A:Title: Genetic variation in the human insulin gene.
 A:Reference number: A94253; MUID:80236313; PMID:6248962
 A:Accession: A94253
 A:Molecule type: DNA
 A:Residues: 1-110 <URL>
 A:Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R:Bell, G.I.; Swain, W.F.; Flicet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.J.
 Nature 282, 525-527, 1979
 A:Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.
 A:Reference number: A93216; MUID:80054779; PMID:503234
 A:Accession: A93216
 A:Molecule type: mRNA
 A:Residues: 1-110 <BRU>
 A:Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R:Sures, I.; Goeddel, D.V.; Gray, A.; Ulrich, A.
 Science 208, 57-59, 1980
 A:Title: Nucleotide sequence of human preproinsulin complementary DNA.
 A:Reference number: A94251; MUID:80147417; PMID:6927840
 A:Accession: A94251
 A:Molecule type: mRNA
 A:Residues: 1-110 <SUR>
 A:Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R:Nicol, D.S.H.W.; Smith, L.F.
 Nature 187, 483-485, 1960
 A:Title: Amino-acid sequence of human insulin.
 A:Reference number: A93144
 A:Accession: A93144
 A:Molecule type: protein
 A:Residues: 25-54;90-110 <NTC>
 R:Over, P.E.; Cho, S.; Peterson, J.D.; Steiner, D.F.
 J Biol. Chem. 246, 1375-1386, 1971
 A:Title: Studies on human proinsulin. Isolation and amino acid sequence of the human part

A:Reference number: A92075; MUID:71116410; PMID:5101771
 A:Accession: A92075
 A:Molecule type: protein
 A:Residues: 57-87 <OYE>
 R:Ko, A.; Smyth, D.G.; Markusen, J.; Sundby, F.
 Eur. J. Biochem. 20, 190-199, 1971
 A:Title: Amino acid sequence of the C-peptide of human proinsulin.
 A:Reference number: A91186; MUID:71257722; PMID:5560404
 A:Accession: A91186
 A:Molecule type: protein
 A:Residues: 57-87 <KOA>
 R:Knudsen, A.M.; Jørgen, C.; Beresni, J.P.; Bolvard, C.; Froguel, P.; Lathrop, M.; Be
 Nature Gene. 4, 305-310, 1993
 A:Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment
 A:Reference number: 158114; MUID:93364428; PMID:8358440
 A:Accession: 158114
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-59,63-110 <RSB>
 A:Cross-references: GB:U15440; NID:g307071; PIDN:AAA59179.1; PID:g307072
 R:Sieber, P.; Kamber, B.; Hartman, A.; Joehl, A.; Riniker, B.; Rittel, W.
 Helv. Chim. Acta 57, 2617-2621, 1974
 A:Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
 A:Reference number: A91636; MUID:75077277; PMID:4443293
 A:Contents: annotation; synthesis
 A:Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identical
 A:Note: article in German with English abstract
 R:Nathani, V.K.
 Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
 A:Title: The synthesis of C-peptide of human proinsulin.
 A:Reference number: A91658; MUID:75040007; PMID:4803504
 A:Contents: annotation; synthesis of residues 57-87
 R:Geiger, R.; Jaeger, G.; Koening, W.
 Chem. Ber. 106, 2347-2352, 1973
 A:Title: Synthesis of the complete sequence of human proinsulin C-peptide and its [Glu
 A:Reference number: A90914
 A:Contents: annotation; synthesis of residues 57-87
 R:Kaufmann, U.B.; Irmlinger, U.C.; Halban, P.A.
 Biochem. J. 310, 869-874, 1995
 A:Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junct
 A:Reference number: S58661; MUID:96031385; PMID:7575420
 A:Contents: annotation; site-directed mutagenesis study of proteolytic processing
 A:Genetics:
 A:Gene: GDB:INS
 A:Cross-references: GDB:119349; OKIM:176730
 A:Map position: 11p15.5-11p15.5
 A:Introns: 63/71
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status experimental <BOH>
 F:25-54,90-110/Product: insulin #status experimental <MAT>
 F:57-87/Domain: connecting C peptide #status experimental <CPE>
 F:90-110/Domain: insulin chain A #status experimental <ACH>
 F:31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 90.8%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 1.6e-23;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLCGSHLVEAALPVCGEGFETPKTKGIVEQCCTISGLYOLENYCN 52
 DB 25 FVNQHLCGSHLVEAALPVCGEGFETPKTKREAEDLVGVLEGGFAGSLPDLAEG 84

QY 31 ----RGIVEQCCTISGLYOLENYCN 52
 DB 85 SLOKRGIVEQCCTISGLYOLENYCN 110

RESULT 9
 B42179
 Insulin precursor - green monkey
 C:Species: Cercopithecus aethiops (green monkey, grivet)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: B42179; A05232; S16494; S22056
R/Species: S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A/Title: Sequences of primate insulin genes support the hypothesis of a slower rate of m
A/Reference number: A42179; MUID:92219953; PMID:1560757
A/Accession: B42179
A/Molecule type: DNA
A/Residues: 1-110 <SEI>
A/Cross-references: UNIPROT:P30407; EMBL:X61092; NID:922808; PIDN:CAA43405.1; PID:922809
A/Note: sequence extracted from NCBI backbone (NCBI:95185, NCBI:95194)
J. Biol. Chem. 247, 4866-4871, 1972
R/Petermann, J.D.; Neimlich, S.; Oyer, P.E.; Steiner, D.F.
A/Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsulin
A/Reference number: A92111; MUID:72258026; PMID:4626369
A/Accession: A05232
A/Molecule type: protein
A/Residues: 57-87 <PRT>
C/Genetics:
A/Intons: 63/1
C/Superfamily: insulin
C/Keywords: hormone; pancreas
F/1-24/Domain: signal sequence #status predicted <SIG>
F/25-54/Domain: insulin chain B #status predicted <BOH>
F/57-87/Domain: insulin chain A #status predicted <MAT>
F/90-110/Domain: connecting peptide #status experimental <CPBP>
F/91-96/43-109/95-100/Disulfide bonds: #status predicted
Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
Qy 1 FVNHLCGSHVLEALYVCGRGFFYPKT----- 30
Db 25 FVNHLCGSHVLEALYVCGRGFFYPKTRAEADLVQVQVLEGGPGAGSLQPLALEG 84
Qy 31 ----RGIVEOCCTSGISLYOLENYCN 52
Db 85 SLQKRGIVEOCCTSGISLYOLENYCN 110
RESULT 10
J00178
Insulin precursor - crab-eating macaque
C/Species: Macaca fascicularis (crab-eating macaque)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C/Accession: J00178
R/Wetzel, W.; Gronenberg, J.; Leinweber, M.; Wengemayer, F.; Winkler, E.L.
Gene 19, 179-183, 1982
A/Title: The nucleotide sequence of cDNA coding for preproinsulin from the primate Maca
A/Reference number: J00178; MUID:83080474; PMID:6184262
A/Accession: J00178
A/Molecule type: mRNA
A/Residues: 1-110 <WET>
A/Cross-references: UNIPROT:P30406; GB:U00336; NID:9342121; PIDN:AAA6849.1; PID:9342122
C/Superfamily: insulin
F/1-24/Domain: signal sequence #status predicted <SIG>
F/25-54/Domain: insulin chain B #status predicted <BOH>
F/57-87/Domain: insulin chain A #status predicted <MAT>
F/90-110/Domain: insulin chain A #status predicted <ACH>
F/91-96/43-109/95-100/Disulfide bonds: #status predicted
Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
Qy 1 FVNHLCGSHVLEALYVCGRGFFYPKT----- 30
Db 25 FVNHLCGSHVLEALYVCGRGFFYPKTRAEADLVQVQVLEGGPGAGSLQPLALEG 84
Qy 31 ----RGIVEOCCTSGISLYOLENYCN 52

Db 85 SLQKRGIVEOCCTSGISLYOLENYCN 110
RESULT 11
Insulin precursor - chimpanzee
C/Species: Pan troglodytes (chimpanzee)
C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A42179; S22058
R/Species: S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A/Title: Sequences of primate insulin genes support the hypothesis of a slower rate of m
A/Reference number: A42179; MUID:92219953; PMID:1560757
A/Accession: A42179
A/Molecule type: DNA
A/Residues: 1-110 <SEI>
A/Cross-references: UNIPROT:P30410; EMBL:X61089; NID:938251; PIDN:CAA43403.1; PID:938252
A/Note: sequence extracted from NCBI backbone (NCBI:95067)
C/Genetics:
A/Intons: 63/1
C/Superfamily: insulin
Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
Qy 1 FVNHLCGSHVLEALYVCGRGFFYPKT----- 30
Db 25 FVNHLCGSHVLEALYVCGRGFFYPKTRAEADLVQVQVLEGGPGAGSLQPLALEG 84
Qy 31 ----RGIVEOCCTSGISLYOLENYCN 52
Db 85 SLQKRGIVEOCCTSGISLYOLENYCN 110
RESULT 12
INCNA
Insulin - Arabian camel (tentative sequence)
C/Species: Camelus dromedarius (Arabian camel)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C/Accession: A92782
R/Dano, W.O.; Baghdad, I.; 16-28, 1972
U. Pac. Med. 16, 16-28, 1972
A/Title: The isolation and characterization of insulin of camel (Camelus dromedarius).
A/Reference number: A92782
A/Accession: A92782
A/Molecule type: protein
A/Residues: 1-30/31-51 <DN>
A/Cross-references: UNIPROT:P01320
C/Superfamily: insulin
C/Keywords: hormone; pancreas
F/1-30/Domain: insulin chain B #status experimental <BOH>
F/31-51/Domain: insulin chain A #status experimental <MAT>
F/57-87/Domain: insulin chain A #status experimental <ACH>
F/90-110/Domain: insulin chain A #status predicted
F/91-96/43-109/95-100/Disulfide bonds: #status predicted
Query Match 89.6%; Score 263.5; DB 1; Length 51;
Best Local Similarity 90.4%; Pred. No. 2.1e-23;
Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Qy 1 FVNHLCGSHVLEALYVCGRGFFYPKTRGIVEOCCTSGISLYOLENYCN 52
Db 1 FVNHLCGSHVLEALYVCGRGFFYPKTRAEADLVQVQVLEGGPGAGSLQPLALEG 84
RESULT 13
INGT
Insulin - goat
C/Species: Capra aegagrus hircus (domestic goat)
C/Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
C/Accession: A01586

R:Smith, L.F.
 Am J Med. 40: 662-666, 1966
 A>Title: Species variation in the amino acid sequence of insulin.
 A:Reference number: A90029; MUID:66160119; PMID:15949593
 A:Accession: A01586
 A:Molecule type: protein
 A:Residues: 1-30/31-51 <SMI>
 A:Cross-references: UNIPROT:P01319
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-30/Domain: insulin chain B #status experimental <ECH>
 F:1-30/31-51/Product: insulin #status experimental <MAT>
 F:31-51/Domain: insulin chain A #status experimental <ACH>
 F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 89.6%; Score 263.5; DB 1; Length 51;
 Best local similarity 90.4%; Pred. No. 2,1e-23;
 Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYOLENYCN 52
 DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKA-GIVEQCAGVCSLYOLENYCN 51

RESULT 14

INMHIS
 Insulin - sei whale
 C:Species: Galeenoptera borealis (sei whale)
 C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
 C:Accession: A01582
 R:Ishihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
 Nature 181, 1468-1469, 1958
 A>Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pepsin.
 A:Reference number: A93142
 A:Accession: A01582
 A:Molecule type: protein
 A:Residues: 1-30/31-51 <ISH>
 A:Cross-references: UNIPROT:P01314
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-30/Domain: insulin chain B #status experimental <ECH>
 F:1-30/31-51/Product: insulin #status experimental <MAT>
 F:31-51/Domain: insulin chain A #status experimental <ACH>
 F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 89.6%; Score 263.5; DB 1; Length 51;
 Best local similarity 92.3%; Pred. No. 2,1e-23;
 Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYOLENYCN 52
 DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKA-GIVEQCAGVCSLYOLENYCN 51

RESULT 15

IPPG
 Insulin precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 16-Jul-1999
 C:Accession: A01583; A94572; S16492; A60835; B60835
 R:Chance, R.E.; Ellis, R.M.; Bromer, W.W.
 Science 161, 165-167, 1968
 A>Title: Porcine proinsulin: characterization and amino acid sequence.
 A:Reference number: A94240; MUID:68266485; PMID:5657063
 A:Accession: A01583

A:Molecule type: protein
 A:Residues: 1-34, Q, 7,36-84 <CHA>
 R:Chance, R.E.
 submitted to the Atlas, July 1970
 A:Reference number: A94572
 A:Accession: A94572
 A:Molecule type: protein
 A:Residues: 1-84 <CH2>

R:Brown, H.; Sanger, F.; Kitai, R.
 Biochem. U. 60, 556-565, 1955
 A>Title: The structure of pig and sheep insulins.
 A:Reference number: A90344
 A:Accession: S16492
 A:Molecule type: protein
 A:Residues: 1-30/31-51 <BRO>

R:Snell, L.; Damgaard, U.
 Horm. Metab. Res. 20, 476-480, 1988
 A>Title: Proinsulin heterogeneity in pigs.
 A:Reference number: A60835; MUID:89032178; PMID:3181865
 A:Accession: A60835
 A:Molecule type: protein
 A:Residues: 33-38,40-62 <SNE>

A>Note: the authors report the characterization of a connecting peptide variant lacking
 A:Accession: B60835
 A:Molecule type: protein
 A:Residues: 33-62 <SN2>

R:Blundell, T.; Dodson, G.; Hodgkin, D.; Mercola, D.
 Adv. Protein Chem. 26, 279-402, 1972

A>Title: Insulin: the structure in the crystal and its reflection in chemistry and bio
 A:Reference number: A90017

C:Superfamily: insulin
 C:Keywords: hormone; pancreas

F:1-30/Domain: insulin chain B #status experimental <ECH>
 F:1-30/64-84/Product: insulin #status experimental <MAT>

F:33-63/Domain: connecting peptide #status experimental <CEPP>
 F:64-84/Domain: insulin chain A #status experimental <ACH>

F:7-70,19-83,69-74/Disulfide bonds: #status experimental

Query Match 89.5%; Score 263; DB 1; Length 84;
 Best local similarity 60.7%; Pred. No. 3,7e-23;
 Matches 51; Conservative 0; Mismatches 1; Indels 32; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYOLENYCN 30

DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYOLENYCN 60

QY 31 --RGIVEQCCTSIQSLYOLENYCN 52

DB 61 QKRGIVEQCCTSIQSLYOLENYCN 84

Search completed: November 2, 2004, 20:22:16
 Job time : 10.5941 secs

This Page Blank (uspio)

Sat Nov 6 18:59:27 2004

us-10-054-873-5.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:20:47 ; Search time 38.952 Seconds
(without alignments)
432.820 Million cell updates/sec

Title: US-10-054-873-5
Perfect score: 294
Sequence: 1 FVNHQHGSHLVKALIVCG.....IYEQCTCTCSLYQLENYCN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgnt2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgnt2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgnt2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgnt2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgnt2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgnt2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgnt2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgnt2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgnt2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgnt2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgnt2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgnt2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
13: /cgnt2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgnt2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgnt2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgnt2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgnt2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgnt2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgnt2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgnt2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	294	100.0	52	US-10-054-873-5
2	294	100.0	107	US-10-054-873-6
3	294	100.0	137	US-10-101-454-39
4	294	100.0	145	US-10-101-454-45
5	294	100.0	146	US-10-101-454-48
6	294	100.0	150	US-10-054-873-7
7	278.5	94.7	51	US-09-858-935B-5
8	278.5	94.7	51	US-10-028-410-3
9	278.5	94.7	51	US-10-444-326-3
10	278.5	94.7	51	US-10-271-869-5
11	278.5	94.7	51	US-10-444-262-3
12	278.5	94.7	51	US-10-444-649-3
13	278.5	94.7	51	US-10-444-701-3

14	275.5	93.7	104	US-10-101-454-15	Sequence 15, Appl
15	275.5	93.7	124	US-09-894-711-18	Sequence 18, Appl
16	275.5	93.7	138	US-09-861-687-19	Sequence 19, Appl
17	275.5	93.7	138	US-10-620-651-19	Sequence 19, Appl
18	275.5	93.7	140	US-10-101-454-33	Sequence 33, Appl
19	275.5	93.7	140	US-10-101-454-42	Sequence 42, Appl
20	273	92.9	50	US-10-066-009A-3	Sequence 36, Appl
21	271.5	92.3	102	US-10-101-454-36	Sequence 3, Appl
22	267	90.8	86	US-09-878-380-1	Sequence 1, Appl
23	267	90.8	86	US-09-858-935B-4	Sequence 2, Appl
24	267	90.8	86	US-10-028-410-2	Sequence 2, Appl
25	267	90.8	86	US-10-054-873-4	Sequence 2, Appl
26	267	90.8	86	US-10-444-326-2	Sequence 2, Appl
27	267	90.8	86	US-10-271-869-4	Sequence 4, Appl
28	267	90.8	86	US-10-444-262-2	Sequence 2, Appl
29	267	90.8	86	US-10-444-649-2	Sequence 2, Appl
30	267	90.8	86	US-10-444-701-2	Sequence 2, Appl
31	267	90.8	96	US-09-947-563-4	Sequence 4, Appl
32	267	90.8	96	US-09-205-658-125	Sequence 125, App
33	267	90.8	110	US-09-815-229-3	Sequence 3, Appl
34	267	90.8	110	US-09-804-409A-9	Sequence 6, Appl
35	267	90.8	110	US-09-963-748C-5	Sequence 125, App
36	267	90.8	110	US-09-963-693-15	Sequence 1, Appl
37	267	90.8	110	US-10-038-686-1	Sequence 1, Appl
38	267	90.8	110	US-10-328-813-2	Sequence 2, Appl
39	267	90.8	110	US-10-383-265-2	Sequence 2, Appl
40	267	90.8	110	US-10-346-563-2	Sequence 2, Appl
41	267	90.8	110	US-10-321-717-2	Sequence 2, Appl
42	267	90.8	110	US-10-411-037-44	Sequence 44, Appl
43	267	90.8	110	US-10-411-026-44	Sequence 44, Appl
44	267	90.8	110	US-10-410-962-44	Sequence 44, Appl
45	267	90.8	110	US-10-411-049-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-10-054-873-5
Sequence 5, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chapterone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 52 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-054-873-5

Query Match 100.0%; Score 294; DB 13; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.7e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKRGIVVEQCCSTISCSLYOLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYPTKRGIVVEQCCSTISCSLYOLENYCN 52

RESULT 2
US-10-054-873-6
Sequence 6, Application US/10054873
Publication No. US20020164712A1

GENERAL INFORMATION:

APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-like Sequence

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/CN98/000052
FILING DATE: 31-Mar-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-054-873-6

Query Match 100.0%; Score 294; DB 13; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKRGIVVEQCCSTISCSLYOLENYCN 52
DB 56 FVNQHLGSHLVEALYLVCGERGFFYPTKRGIVVEQCCSTISCSLYOLENYCN 107

RESULT 3
US-10-101-454-39
Sequence 39, Application US/10101454

Publication No. US20040110664A1

GENERAL INFORMATION:

APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, Ib
Andersen, Asger Sloth
Markussen, Jan

TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:

ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/400,256
FILING DATE: 03-Mar-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-10-101-454-39

Query Match 100.0%; Score 294; DB 16; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKRGIVVEQCCSTISCSLYOLENYCN 52
DB 86 FVNQHLGSHLVEALYLVCGERGFFYPTKRGIVVEQCCSTISCSLYOLENYCN 137

RESULT 4
US-10-101-454-45
Sequence 45, Application US/10101454
Publication No. US20040110664A1

GENERAL INFORMATION:

APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, Ib
Andersen, Asger Sloth
Markussen, Jan

TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-10-101-454-45

Query Match 100.0%; Score 294; DB 16; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.6e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
US-10-101-454-48
Sequence 48, Application US/10101454
Publication No. US20040110664A1
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, Ib
Andersen, Asger Sloth
Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-101-454-48

Query Match 100.0%; Score 294; DB 16; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.6e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISCLYQLENYCN 52
Db 95 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISCLYQLENYCN 146

RESULT 6
US-10-054-873-7
Sequence 7, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chapterone-Like Sequence

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match 100.0%; Score 294; DB 13; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.7e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISCLYQLENYCN 52
Db 99 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISCLYQLENYCN 150

RESULT 7

```
US-09-858-935B-5
; Sequence 5, Application US/09858935B
; Publication No. US2003069177A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Lowman, Henry B.
; TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
; FILE REFERENCE: P1794R1
; CURRENT APPLICATION NUMBER: US/09/858,935B
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/248,985
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/204,490
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 5
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-935B-5

Query Match          94.7%; Score 278.5; DB 10; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 51

RESULT 8
US-10-028-410-3
; Sequence 3, Application US/10028410
; Publication No. US20020160955A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/10/028,410
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/09/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-410-3

Query Match          94.7%; Score 278.5; DB 13; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 51

RESULT 9
US-10-444-326-3
; Sequence 3, Application US/10444326
; Publication No. US20030191065A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,326
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/723,866
; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-326-3

; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-326-3

; PRIOR APPLICATION NUMBER: US/10271869
; Publication No. US20030211992A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Lowman, Henry B.
; TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
; FILE REFERENCE: P1794R1
; CURRENT APPLICATION NUMBER: US/10/271,869
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US/09/858,935
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/248,985
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/204,490
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 5
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-271-869-5

Query Match          94.7%; Score 278.5; DB 14; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 51

RESULT 10
US-10-271-869-5
; Sequence 5, Application US/10271869
; Publication No. US20030211992A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Lowman, Henry B.
; TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
; FILE REFERENCE: P1794R1
; CURRENT APPLICATION NUMBER: US/10/271,869
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US/09/858,935
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/248,985
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/204,490
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 5
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-271-869-5

Query Match          94.7%; Score 278.5; DB 14; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIGSLYLENYCN 51

RESULT 11
US-10-444-262-3
; Sequence 3, Application US/10444262
; Publication No. US20040023883A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,262
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/724,478
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-262-3
```

Query Match 94.7%; Score 278.5; DB 15; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 51

RESULT 12
US-10-444-649-3
; Sequence 3, Application US/10444649
; Publication No. US20040033951A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,649
; PRIOR FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/724,479
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-649-3

Query Match 94.7%; Score 278.5; DB 15; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 51

RESULT 13
US-10-444-701-3
; Sequence 3, Application US/10444701
; Publication No. US20040033952A1
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,701
; PRIOR FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/723,666
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-701-3

Query Match 94.7%; Score 278.5; DB 15; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.1e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 51

RESULT 14

US-10-101-454-15
; Sequence 15, Application US/10101454
; Publication No. US20040110664A1
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Joransen, Ib
; APPLICANT: Andersen, Aaseer Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Parentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/101,454
; FILING DATE: 20-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Landis, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-101-454-15

Query Match 93.7%; Score 275.5; DB 16; Length 104;
Best Local Similarity 90.9%; Pred. No. 1.9e-25;
Matches 50; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGISLVQLENYCN 52
DB 50 FVNQHLGSHLVEALYLVCGERGFFYTPKSDAKGIVECCCTSGISLVQLENYCN 104

RESULT 15
US-09-894-711-18
; Sequence 18, Application US/09894711
; Patent No. US20020137144A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Borglum
; APPLICANT: Ludvigsen, Svend
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: Insulin precursor analogues having improved fermentation
; TITLE OF INVENTION: Yield in yeast
; FILE REFERENCE: 6148.400-US
; CURRENT APPLICATION NUMBER: US/09/894,711
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: PA 2000 00443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PA 1999 01869
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/211,081

PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: 60/181,450
PRIOR FILING DATE: 2000-02-10
PRIOR APPLICATION NUMBER: 09/740,359
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 18
LENGTH: 124
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-894-711-18

Query Match 93.7%; Score 275.5; DB 9; Length 124;
Best Local Similarity 94.3%; Pred. No. 2,3e-25;
Matches 50; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
Cy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPK-TRGIVEOCCCTSIQSLYQLENYCN 52
Db 72 FVNQHLCGSHLVEALYLVCGERGFFYTPKAKGIIVEOCCCTSIQSLYQLENYCN 124

Search completed: November 2, 2004, 20:59:22
Job time : 38.952 secs

Sat Nov 6 18:59:28 2004

us-10-054-873-5.rup

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:48:36; Search time 50.365 Seconds
(without alignments)
594.006 Million cell updates/sec

Title: US-10-054-873-5

Perfect score: 294

Sequence: 1 FVNOHLCGSHLVEALYLVCG.....IVEQCCTSIICSLYQENYCN 52

Scoring table:

BLOSUM62
Gapop 10.0, Gapept 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_sprot; 2: uniprot_trembl; 3: uniprot_swissprot

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	273.5	93.0	51	INS_BALPH	P01312 balaenopter
2	273.5	93.0	51	INS_ELEMA	P01316 elephas max
3	273	92.9	96	Q7M0U6	Q7M0U6 bacillus br
4	271.5	92.3	51	Q7M0G1	Q7M0G1 cricetidae
5	268.5	91.3	51	INS_ACOCA	P01324 acomyx canl
6	267.5	91.0	51	Q7M217	Q7M217 canavalia e
7	267	90.8	110	INS_CERAE	P30407 cercopithec
8	267	90.8	110	INS_HUMAN	P01308 homo sapien
9	267	90.8	110	INS_PANTR	P30406 macaca fasc
10	267	90.8	110	INS_PANTR	P30410 pan troglod
11	267	90.8	110	INS_PONY	Q6YK33 gorilla gor
12	267	90.8	110	Q6YK33	Q6YK33 gorilla gor
13	267	90.8	110	AAP35454	AAP35454 homo sapi
14	267	90.8	110	AAN06935	AAN06935 gorilla g
15	267	90.8	110	AAN39451	AAN39451 homo sapi
16	263.5	89.6	51	INS_BALBO	P01314 balaenopter
17	263.5	89.6	51	INS_CAMPR	P01320 camelus dro
18	263.5	89.6	51	INS_CAMPR	P01315 capra hircu
19	263	89.5	108	INS_PIG	P01313 sus scrofa
20	263	89.5	108	AAQ00952	AAQ00952 sus scrof
21	263	89.5	108	AAQ00954	AAQ00954 sus scrof
22	263	89.5	108	AAQ00957	AAQ00957 sus scrof
23	263	89.5	108	AAQ00960	AAQ00960 sus scrof
24	263	89.5	108	AAQ00963	AAQ00963 sus scrof
25	263	89.5	108	AAQ00966	AAQ00966 sus scrof
26	263	89.5	108	AAQ00969	AAQ00969 sus scrof
27	263	89.5	108	AAQ00972	AAQ00972 sus scrof
28	263	89.5	108	AAQ00975	AAQ00975 sus scrof
29	263	89.5	108	AAQ00978	AAQ00978 sus scrof
30	263	89.5	108	AAQ00981	AAQ00981 sus scrof
31	263	89.5	108	AAQ00983	AAQ00983 sus scrof

ALIGNMENTS

RESULT 1	ID	INS_BALPH	STANDARD	PRT	51 AA.
AC	P01312				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	05-JUL-2004 (Rel. 44, Last annotation update)				
DE	Insulin.				
GN	Name=INS;				
OS	Balaenoptera physalus (Finback whale) (Common rorqual), and				
OS	Physeter catodon (Sperm whale) (Physeter macrocephalus).				
OC	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;				
OC	Balaenopteridae; Balaenoptera.				
CX	NCBI_TaxID=9770, 9755;				
RN	[1]				
RN	PARTIAL SEQUENCE.				
RC	SPECIES=B.physalus;				
RA	Hama H., Tiltan K., Sakaki S., Narita K.;				
RT	"The amino acid sequence in fin-whale insulin.";				
RL	J. Biochem. 56:285-293(1964).				
RN	[2]				
RP	SEQUENCE.				
RC	SPECIES=P.catodon;				
RA	Harris U.I., Sanger F., Naughton M.A.;				
RT	"Species differences in insulin.";				
RL	Arch. Biochem. Biophys. 65:427-438(1956).				
CC	-1- FUNCTION: Insulin decreases blood glucose concentration. It				
CC	increases cell permeability to monosaccharides, amino acids and				
CC	fatty acids. It accelerates glycolysis, the pentose phosphate				
CC	cycle, and glycogen synthesis in liver.				
CC	-1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two				
CC	disulfide bonds.				
CC	-1- STRUCTURAL LOCATION: Secreted.				
CC	-1- SIMILARITY: Belongs to the insulin family.				
DR	PIR; A91918; INWHF.				
DR	PIR; A91918; INWHF.				
DR	HSP; P01317; IAPF.				
DR	InterPro: PR004825; Ins/IGF/relax.				
DR	PRINTS: PR00277; INSULIN.				
DR	SMART: SM00078; IIGF. 1.				
DR	PROSITE: PS00252; INSULIN; 1.				
KW	Direct protein sequencing; Glucose metabolism; Hormone;				
KW	Insulin family.				
FT	CHAIN 1				
FT	NON_CONS 30				
FT	CHAIN 31				
FT	NON_CONS 51				
FT	CHAIN				

AAQ00985 sus scrof	AAQ00987 sus scrof	AAQ00990 sus scrof	P01311 oryctolagus	Q91X13 spermophilu	P06306 felis silve	P01321 canis faml	P01313 cricetus	P01317 bos taurus	Q62587 psammomyx o	P18109 didelphis m	P01318 ovis aries	P01310 equus cabal	P01327 chinchilla
--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	-----------------	-------------------	--------------------	--------------------	-------------------	--------------------	-------------------

```

FT DISULFID 7 37 Interchain.
FT DISULFID 19 50 Interchain.
FT DISULFID 36 41 Interchain.
SQ SEQUENCE 51 AA; 5768 MW; 9007B514691A7CDD CRC64;

Query Match 93.0%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 4e-26;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 51

RESULT 2
INS_ELEMA STANDARD; PRT; 51 AA.
ID INS_ELEMA
AC P01316;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin.
GN Name=INS;
OS Elephas maximus (Indian elephant).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
OC NCBI_Taxid=9783;
RN [1]
RP SEQUENCE.
RX MEDLINE=66160119; PubMed=5949593;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin.";
RL Am. J. Med. 40:662-666(1966).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: The species of elephant is not given, but it is
CC most probably the Indian elephant (Elephas maximus).
CC -1- SIMILARITY: Belongs to the insulin family.
DR HSSP; P01308; 1A10.
DR InterPro; IPR004825; Ins/IGF/relax.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IIGF.1.
DR PROSITE; PS00262; INSULIN.1
KW Direct protein sequencing; Glucose metabolism; Hormone;
KW Insulin family.
FT CHAIN 1 30 Insulin B chain.
FT NON_CONS 30 31
FT CHAIN 31 51 Insulin A chain.
FT DISULFID 7 37 Interchain.
FT DISULFID 19 50 Interchain.
FT DISULFID 36 41 Interchain.
SQ SEQUENCE 51 AA; 5752 MW; 9007B50CDB457DED CRC64;

Query Match 93.0%; Score 273.5; DB 1; Length 51;
Best Local Similarity 94.2%; Pred. No. 4e-26;
Matches 49; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 51

RESULT 3
OY 07M06 PRELIMINARY; PRT; 96 AA.
AC 07M06;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)

```

```

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Epidermal growth factor/single chain insulin fusion protein
DE (fragment).
OS Bacillus brevis (Brevibacillus brevis).
OC Bacteria; Firmicutes; Bacillales; Brevibacillaceae; Brevibacillus.
OC NCBI_Taxid=1393;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20335834; PubMed=10879487;
RA Koh M., Hanagata H., Eblus S., Morihara K., Takagi H.;
RT "Use of Bacillus brevis for synthesis and secretion of Des-B30 single-
RT chain human insulin precursor.";
RL Biosci. Biotechnol. Biochem. 64:1079-1081(2000).
DR PIR; PCT082; PCT082.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0007582; P:physiological process; IEA.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin.1.
DR PRINTS; PR00277; INSULINB.
DR PROSITE; PS00262; INSULIN.1.
FT NON_TER 1 1
FT NON_TER 96 96
SQ SEQUENCE 96 AA; 10473 MW; 4505D710C289092A CRC64;

Query Match 92.9%; Score 273; DB 2; Length 96;
Best Local Similarity 96.2%; Pred. No. 8.5e-26;
Matches 50; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

OY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 52
DB 47 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 96

RESULT 4
OY 07M06 PRELIMINARY; PRT; 51 AA.
ID 07M06;
AC 07M06;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Insulin.
OS Cricetidae sp. (Hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae.
OC NCBI_Taxid=36483;
RN [1]
RP SEQUENCE.
RA Neelson F.A., Delcher H.K., Steinman H., Lebovitz H.E.;
RT "Structure of hamster insulin: comparison with a tumor insulin.";
RT Fed. Proc. 32:300-300(1973).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the insulin family.
DR PIR; A91456; A91456.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0007582; P:physiological process; IEA.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin.1.
DR PRINTS; PR00277; INSULINB.
DR PROSITE; PS00262; INSULIN.1.
KW Insulin family.
SQ SEQUENCE 51 AA; 5768 MW; 9006B6469047D3D CRC64;

Query Match 92.3%; Score 271.5; DB 2; Length 51;
Best Local Similarity 94.2%; Pred. No. 7e-26;
Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

OY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRIGIVEOCCSTISLYOLENYCN 51

```

```

RESULT 5
INS_ACOCA STANDARD: PRT; 51 AA.
ID INS_ACOCA
AC P01324;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin.
GN Name-INS;
OS Acomys cahirinus (Egyptian spiny mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Acomys.
OX NCBI_TaxID=10068;
RN (1)
RP PRELIMINARY SEQUENCE.
RX MEDLINE=72189454; PubMed=5028210;
RA Buenzli H.P., Humbel R.E.;
RT "Isolation and partial structural analysis of insulin from mouse (Mus musculus) and spiny mouse (Acomys cahirinus).";
RL Hoppe-Seyler's Z. Physiol. Chem. 353:444-450(1972).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
DR PIR, A01591; IMWSP.
DR HSSP, P01308; IEV6.
DR InterPro: IPR004825; Ins/IGF/relax.
DR PRINTS; PR00277; INSULINB.
DR SMART; SMO0078; IIGF.1.
DR PROSITE; PS00262; INSULIN; 1.
KW Direct protein sequencing; Glucose metabolism; Hormone;
KW Insulin family.
KM CHAIN 1 30 Insulin B chain.
FT NON_CONS 1 31
FT CHAIN 31 51 Insulin A chain.
FT DISULFID 7 37 Interchain (By similarity).
FT DISULFID 19 50 Interchain (By similarity).
FT DISULFID 36 41 By similarity.
SQ SEQUENCE 51 AA; 5768 MW; 9928DB8629047D3D CRC64;

Query Match 91.0%; Score 267.5; DB 2; Length 51;
Best Local Similarity 92.3%; Pred. No. 2.2e-25;
Matches 48; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

1 FVNOHLCGSHLVEALYLVCGERGFFYPTKTRGIYEQCTSIICSLYOLENYCN 52
1 FVNOHLCGSHLVEALYLVCGERGFFYPTKRS-GIVDQCCTSIICSLYOLENYCN 51

```

```

RESULT 7
INS_CERAF STANDARD: PRT; 110 AA.
ID INS_CERAF
AC P30407; P01309;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin precursor.
GN Name-INS;
OS Cercopithecus aethiops (Green monkey) (Givet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92219953; PubMed=1560757;
RA Saino S., Bell G.I., Li W.;
RT "Sequences of primate insulin genes support the hypothesis of a slower rate of molecular evolution in humans and apes than in monkeys.";
RT Mol. Biol. Evol. 9:193-203(1992).
RL [2]
RP SEQUENCE OF 57-87.
RX MEDLINE=72258016; PubMed=4626369;
RA Peterson J.D., Nehrlich S., Oyer P.E., Steiner D.F.;
RT "Determination of the amino acid sequence of the monkey, sheep, and rat dog proinsulin C-peptides by a semi-ntro Edman degradation procedure.";
RL J. Biol. Chem. 247:4866-4871(1972).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).
CC EMBL; X61092; CAA43405.1;
CC PIR; B42179; B42179.
CC HSSP; P01308; IAI0.

```


RN [14]
 RP SYNTHESIS OF 84-87. PubMed=469853;
 RX MEDLINE=73161261; PubMed=469853;
 RA Geiger R., Jaeger G., Keong M., Treuth G.;
 RT "Synthesis of peptides with the properties of human proinsulin C
 RT peptides (hC peptide). I. Scheme for the synthesis and preparation of
 RT the sequence 28-31 of human proinsulin C peptide.";
 RL Chem. Ber. 106:188-192(1973).
 RN [15]
 RP VARIANT LOS ANGELES SER-48.
 RX MEDLINE=84016053; PubMed=6312455;
 RA Hanada M., Chan S.J., Kwok S.C.W., Rubenstein A.H., Steiner D.F.;
 RT "Studies on mutant human insulin genes: identification and sequence
 RT analysis of a gene encoding Iser241insulin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
 RN [16]
 RP VARIANTS LOS ANGELES SER-48 AND CHICAGO LEU-49.
 RX MEDLINE=84170233; PubMed=6424111;
 RA Shoelson S., Fickova M., Hanada M., Nahum A., Musco G., Kaiser E.T.,
 RT Rubenstein A.H., Tager H.;
 RT "Identification of a mutant human insulin predicted to contain a
 RT serine-for-phenylalanine substitution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
 RN [17]
 RP VARIANT PROVIDENCE ASP-34.
 RX MEDLINE=87175640; PubMed=3470784;
 RA Chan S.J., Seino S., Gruppuso P.A., Schwartz R., Steiner D.F.;
 RT "A mutation in the B chain coding region is associated with impaired
 RT proinsulin conversion in a family with hyperproinsulinemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
 RN [18]
 RP VARIANT WAKAYAMA LEU-92.
 RX MEDLINE=87058122; PubMed=3537011;
 RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
 RT "Structurally abnormal insulin in a diabetic patient. Characterization
 RT of the mutant insulin A3 (Val--Ileu) isolated from the pancreas.";
 RL J. Clin. Invest. 78:1666-1672(1986).
 RN [19]
 RP VARIANT HIS-89.
 RX MEDLINE=903117021; PubMed=2196279;
 RA Barbetti F., Raben N., Kadowaki T., Cama A., Accili D., Gabbay K.H.,
 RT Merenich J.A., Taylor S.I., Roth J.;
 RT "Two unrelated patients with familial hyperproinsulinemia due to a
 RT mutation substituting histidine for arginine at position 65 in the
 RT proinsulin molecule: identification of the mutation by direct
 RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
 RT chain reaction.";
 RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
 RN [20]
 RP VARIANT HIS-89.
 RX MEDLINE=85261996; PubMed=4019786;
 RA Shibaaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.;
 RT "Posttranslational cleavage of proinsulin is blocked by a point
 RT mutation in familial hyperproinsulinemia.";
 RL J. Clin. Invest. 89:1902-1907(1992).
 RN [21]
 RP VARIANT KYOTO LEU-89.
 RX MEDLINE=92291307; PubMed=1601997;
 RA Yano H., Kitaro N., Morimoto M., Polonsky K.S., Imura H., Seino Y.;
 RT "A novel point mutation in the human insulin gene giving rise to
 RT hyperproinsulinemia (proinsulin Kyoto).";
 RL J. Clin. Invest. 89:1902-1907(1992).
 RN [22]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91104966; PubMed=2271664;
 RA Hua Q.-X., Weiss M.A.;
 RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
 RT assignment of a des-pentapeptide analogue and comparison with crystal
 RT structure.";
 RL Biochemistry 29:10545-10555(1990).
 RN [23]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91242467; PubMed=2036420;

RA Hua Q.-X., Weiss M.A.;
 RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
 RT insulin: sequential resonance assignment and implications for protein
 RT dynamics and receptor recognition.";
 RL Biochemistry 30:5505-5515(1991).
 RN [24]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91265527; PubMed=1646635;
 RA Hua Q.-X., Weiss M.A.;
 RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-
 RT specific resonance assignments and effects of solvent composition.";
 RL Biochim. Biophys. Acta 1078:110-110(1991).
 Qy Query Match 90.8%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 5,3e-25;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
 Db 1 FVNHQCGSHVLEALVNGSGFFPTKT----- 30
 25 FVNHQCGSHVLEALVNGSGFFPTKTREARLDQGVYELCGSGAGSLQPLALEG 84
 Qy 31 ----RGVEQCCTSTCSLYOLENYCN 52
 85 SLOKRGVEQCCTSTCSLYOLENYCN 110
 RESULT 9
 INS_MACFA
 ID INS_MACFA STANDARD; PRT; 110 AA.
 AC P30406; P01309;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name:INS;
 OS Macaca fascicularis (Cray eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83080474; PubMed=6184262;
 RA Wexler M., Groneberg J., Leineweber M., Wengenmayer F.,
 RT Winkler E.-L.;
 RT "The nucleotide sequence of cDNA coding for preproinsulin from the
 RT primate Macaca fascicularis.";
 RL Gene 19:179-183(1982).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; J00336; AAA36849.1; .
 DR PIR; J00178; J00178.
 DR HSP; P01308; IAI0.
 DR Interpro; IPRO04825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULIN.
 DR SMART; SM00078; IIGF.1
 DR PROSITE; PS00262; INSULIN; 1.

KM Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54
 FT PROPEP 57 87
 FT CHAIN 90 110
 FT DISULFID 31 96
 FT DISULFID 43 109
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA; 11991 MW; 83CGE33A80A420F9 CRC64;
 Query Match 90.8%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 5.3e-25;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
 QY 1 FVNHGSHLVKLVKGVGGRGFYTPKT----- 30
 DB 25 FVNHGSHLVKLVKGVGGRGFYTPKTRRREADLPQGVYELGGPGAGSLQPLALEG 84
 QY 31 ----RGIVEOCCTSGSLVQLENYCN 52
 DB 85 SLQKRGIVEOCCTSGSLVQLENYCN 110
 RESULT 10
 INS_PANTR STANDARD; PRT; 110 AA.
 AC P30410.
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name:INS;
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92219953; PubMed=1560757;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X61089; CAA43403.1; -
 DR EMBL; AY137497; AAN06933.1; -
 DR PIR; A42179; A42179.
 DR HSSP; P01308; IAI0.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.

DR PROSITE; PS00262; INSULIN; 1.
 KM Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54
 FT PROPEP 57 87
 FT CHAIN 90 110
 FT DISULFID 31 96
 FT DISULFID 43 109
 FT DISULFID 95 100
 FT HELIX 32 43
 FT TURN 60 61
 FT HELIX 65 71
 FT HELIX 76 82
 FT HELIX 82 82
 SQ SEQUENCE 110 AA; 12025 MW; 41EB8DF79837CEFS CRC64;
 Query Match 90.8%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 5.3e-25;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
 QY 1 FVNHGSHLVKLVKGVGGRGFYTPKT----- 30
 DB 25 FVNHGSHLVKLVKGVGGRGFYTPKTRRREADLPQGVYELGGPGAGSLQPLALEG 84
 QY 31 ----RGIVEOCCTSGSLVQLENYCN 52
 DB 85 SLQKRGIVEOCCTSGSLVQLENYCN 110
 RESULT 11
 INS_PONPY STANDARD; PRT; 110 AA.
 AC Q8HXV2;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name:INS;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878; DOI=10.1101/gr.948003;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AY137503; AAN06937.1; -
 DR HSSP; P01308; IAI0.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR SMART; SMC0078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KM Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54

FT CHAIN 25 54 Insulin B chain.
FT PROPEP 57 87 C peptide.
FT CHAIN 90 110 Insulin A chain.
FT DISULFID 31 96 Interchain (By similarity).
FT DISULFID 43 109 Interchain (By similarity).
FT DISULFID 95 100 By similarity.
SQ SEQUENCE 110 AA; 12038 MW; 22D2B32B94F520F8 CRC64;

Query Match 90.8%; Score 267; DB 1; Length 110;
Best Local Similarity 60.5%; Pred. No. 5.3e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTR----- 30
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRREADLQVGVELGGPGAGSLQPLALEG 84
QY 31 ----RGIVECCCTSICTSLYOLENYCN 52
DB 85 SLQKRGIVECCCTSICTSLYOLENYCN 110

RESULT 12
Q6YK33 PRELIMINARY; PRT; 110 AA.

AC Q6YK33; 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Insulin.
GN Name=INS;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22833521; PubMed=12952878;
RA Stead J.D., Hurles M.E., Jeffreys A.J.;
RT "Global haplotype diversity in the human insulin gene region."
RL Genome Res. 13:2101-2111(2003).
RL [2]
RP SEQUENCE FROM N.A.
RA Stead J.D.H., Jeffreys A.J.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the insulin family.
CC -1- SIMILARITY: Belongs to the insulin family.
DB EMBL; AY137500; AAN06935.1;
DR InterPro; IPR004825; Ins/IGF/relax.
DR InterPro; IPR00334; Molnuc_ins.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULIN.
DR ProDom; PD015667; Molnuc_ins; 1.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family.
SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 5.3e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTR----- 30
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRREADLQVGVELGGPGAGSLQPLALEG 84
QY 31 ----RGIVECCCTSICTSLYOLENYCN 52
DB 85 SLQKRGIVECCCTSICTSLYOLENYCN 110

RESULT 13
AAP35454 PRELIMINARY; PRT; 110 AA.

AC AAP35454; 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Insulin.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koudinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Pheasant M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator(TM) System Donor
vector."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT006808; AAP35454.1;
SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 5.3e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTR----- 30
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRREADLQVGVELGGPGAGSLQPLALEG 84
QY 31 ----RGIVECCCTSICTSLYOLENYCN 52
DB 85 SLQKRGIVECCCTSICTSLYOLENYCN 110

RESULT 14
AAN06935 PRELIMINARY; PRT; 110 AA.

AC AAN06935; 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Insulin.
GN INS.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22833521; PubMed=12952878;
RA Stead J.D., Hurles M.E., Jeffreys A.J.;
RT "Global haplotype diversity in the human insulin gene region."
RL Genome Res. 13:2101-2111(2003).
RL [2]
RP SEQUENCE FROM N.A.
RA Stead J.D.H., Jeffreys A.J.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137500; AAN06935.1;
SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 5.3e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTR----- 30
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRREADLQVGVELGGPGAGSLQPLALEG 84
QY 31 ----RGIVECCCTSICTSLYOLENYCN 52
DB 85 SLQKRGIVECCCTSICTSLYOLENYCN 110

RESULT 15

AAN39451
 ID AAN39451 PRELIMINARY; PRT; 110 AA.
 AC AAN39451;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, last annotation update)
 DE Insulin.
 GN INS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primata; Catarrhini; Homiidae; Homo.
 OX NCBL_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878;
 RA Stead J.D., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stead J.D.H., Jeffreys A.J.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY138590; AAN39451.1; -
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 90.8%; Score 267; DB 2; Length 110;
 Best local similarity 60.5%; Pred. No. 5,3e-25;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNHGCGSHVEALYVCGERGFPYTPKT----- 30
 DB 25 FVNHGCGSHVEALYVCGERGFPYTPKTRREABDLQVGYELGGPGAGSLQPLALEG 84
 QY 31 ----RGIVEOCTSTICSLYOLENYCN 52
 DB 85 SLQKRGIVEOCTSTICSLYOLENYCN 110

Search completed: November 2, 2004, 20:20:33
 Job time : 50.369 secs